*Optimal Retirement in a General Market Environment* 

# Zhou Yang, Hyeng Keun Koo & Yong Hyun Shin

**Applied Mathematics & Optimization** 

ISSN 0095-4616

Appl Math Optim DOI 10.1007/s00245-020-09671-6





Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to selfarchive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".





## **Optimal Retirement in a General Market Environment**

Zhou Yang<sup>1</sup> · Hyeng Keun Koo<sup>2</sup> · Yong Hyun Shin<sup>3</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

We study an optimal retirement, consumption/portfolio selection problem of an economic agent in a non-Markovian environment. We show that under a suitable condition the optimal retirement decision is to retire when the individual's wealth reaches a threshold level. We express the value and the optimal strategy by using the strong solution of the backward stochastic partial differential variational inequality (BSPDVI) associated with the dual problem. We derive properties of the value function and the optimal strategy by analyzing the strong solution and the free boundary of the BSPDVI. We also make a methodological contribution by proposing an approach to investigate properties of the strong solution and the stochastic free boundary of BSPDVI by combining a probabilistic method and the theory of backward stochastic partial differential equations (BSPDEs).

**Keywords** Backward stochastic partial differential variational inequality  $\cdot$  Stochastic free boundary problem  $\cdot$  Early retirement  $\cdot$  Portfolio selection  $\cdot$  Consumption  $\cdot$  Leisure  $\cdot$  Non-Markovian market environment

## Mathematics Subject Classification 91G10

Hyeng Keun Koo hkoo@ajou.ac.kr

> Zhou Yang yangzhou@scnu.edu.cn

Yong Hyun Shin yhshin@sookmyung.ac.kr

<sup>3</sup> Department of Mathematics, Sookmyung Women's University, Seoul 04310, Republic of Korea

Yang is partially supported by NNSF of China [Grant No. 11771158, 11801091] and Guangdong Basic and Applied Basic Research Foundation [Grant No. 2019A1515011338]. Koo gratefully acknowledges support by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea [Grant No. NRF-2019S1A5A2A03054249]. Shin is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) [Grant No. NRF-2019R1H1A2079177].

<sup>&</sup>lt;sup>1</sup> School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China

<sup>&</sup>lt;sup>2</sup> Department of Financial Engineering, Ajou University, Suwon 16499, Republic of Korea

#### JEL Classification D11 · D12 · D91 · E21 · G11

## **1** Introduction

There is a surging interest in an individual's optimization problem which combines the choice of voluntary retirement and the choices of optimal consumption and an optimal investment portfolio (Choi and Shim [4], Fahri and Panageas [11], Choi, Shim and Shin [5], Dybvig and Liu [9], Yang and Koo [30]). Most of the research in this field, however, has been focused on deriving and studying optimal policies in an environment where the investment opportunity is constant, and thus Markovian, and composed of two assets, a riskless bond and a risky stock. The financial market we observe does not satisfy these simplifying assumptions. The interest rate changes stochastically and there exist a large number of assets whose returns exhibit covariances which change randomly over time (see e.g., Ball and Torous [1]). Furthermore, the changes do not necessarily satisfy the Markovian assumption. For example, the Heath–Jarrow–Morton model (Heath et al. [13]), one of the most commonly used models of the term structure of interest rates, is typically implemented with parameters which do not allow a simple transformation to a Markovian model. Another example is provided by the momentum effect in asset returns (see e.g., Li and Liu [22]).

In this paper we study the problem in a general financial market where there are many assets and the returns of the assets are not necessarily Markovian. We employ the flexible labor supply model of Bodie et al. [3] and Choi et al. [5]. We make the following contributions in this paper. Firstly, we provide a complete theoretical treatment of the optimal consumption and investment problem with an early retirement option in a non-Markovian market environment with a verification theorem. Specifically, we show that the optimal retirement decision is to retire when the individual's wealth reaches a threshold level under a suitable condition. We express the value and the optimal strategy by using the strong solution to the backward stochastic partial differential variational inequality (BSPDVI) associated with the dual problem. Secondly, we derive properties of the value function and the optimal strategy by analyzing the strong solution and the free boundary of the BSPDVI. Thirdly, we make a methodological contribution by proposing an approach to investigate properties of the strong solution and the stochastic free boundary of BSPDVI.

Three common methods have been used to analyze the properties of the value function and optimal strategy of a pure optimal control problem (without a choice of a stopping time): first, the martingale method with a dual transformation (see e.g., Cox and Huang [6], Karatzas et al. [18], Karatzas and Shreve [19]), second, the transformation of the associated Hamilton-Jacobi-Bellman (HJB) equation into a linear partial differential equation (PDE) through the Legendre transformation (see e.g., Karatzas et al. [17]), third, the stochastic maximum principle (see e.g., Yong and Zhou [32]). As explained in Yang and Koo [30], however, all the methods cannot be directly applied to our problem, since the state equations before and after the retirement date are not the same, and the dual transformations and the Legendre transformations in the two stages are different. Moreover, the stopping time is unknown and interacts with the optimal control, and the associated HJB equation of the problem is a backward stochastic partial differential equation (BSPDE), not a PDE.

For an optimal stopping problem, three methods have been commonly used: the martingale and probabilistic method (see e.g., Karatzas and Shreve [19], Karatzas and Wang [20], Nutz and Zhang [24], Peskir and Shiryaev [25]), the backward stochastic differential equation method (see e.g., El Karoui et al. [10]) and the PDE method with analysis of the variational inequality (VI) associated with the problem (see e.g., Frideman [12], Bensoussan and Frideman [2]). But it is difficult to discover the properties of the value function and the optimal strategy only by the martingale method or by the probabilistic method, since the optimal stopping time discovered with these methods is abstract. The PDE method attempts to identify the optimal stopping boundary for a Markovian optimal stopping problem, but cannot be directly applied to our problem, because the problem is a non-Markovian optimal stopping problem including portfolio selection, and the HJB equation associated with the problem is a fully nonlinear BSPDVI. Due to the highly non-linear nature, it is difficult to discover the characteristics of the HJB equation even in the simple Markovian case when the investment opportunity is constant and comprised of two assets (see Yang and Koo [30]).

Yang and Koo [30] study a similar optimal retirement and portfolio selection problem in a Markovian market environment. They propose an approach involving successive transformations and recast the problem into that of a variational inequality satisfied by the dual value function. They derive properties of the value function and the optimal strategy by analyzing the VI and the dual value function, by applying the theory of PDEs. Since the HJB equation associated with our problem is a BSPDE, the methods in Yang and Koo [30] are not directly applicable to the non-Markovian problem.

In this paper we propose a new approach to the optimal retirement and portfolio choice problem in a non-Markovian environment. First, we recast the problem into that of a linear BSPDVI by successive transformations similar to those in Yang and Koo [30], in an attempt to overcome the difficulty associated with the combination of control and optimal stopping problems and the non-linearity of the HJB equation of the primal problem. Next, we provide a verification theorem which provides a connection between the solution to the BSPDVI and the value of the original problem. The theorem also provides a characterization of an optimal strategy by using the solution to the BSPDVI. We then apply the theory of quasilinear BSPDVIs developed in Yang and Tang [31] to show the existence and uniqueness of a strong solution. We verify the assumptions of the verification theorem by combining analysis of the BSPDVI and a probabilistic approach. Finally, we discover properties of the value and the optimal retirement boundary by means of the comparison theory for BSPDVIs. It is worth mentioning that the optimal retirement boundary is random and has no closed form, and hence, analysis of its properties is substantially challenging.

Despite the remarkable achievement in the theory of backward stochastic partial differential equations (BSPDEs, see Hu and Peng [14], Hu et al. [15], Du and Tang [8], and Qiu and Wei [26], etc.),<sup>1</sup> it is still not as mature as that of partial differential equations (PDEs), most of the existing results are about the existence and uniqueness

<sup>&</sup>lt;sup>1</sup> See also Yang and Tang [31] and Koo et al. [21] for more development on BSPDEs and their applications.

of the strong or weak solutions, and investigation of properties of solutions are yet to come. We overcome the difficulty by combining a probabilistic method and the theory of BSPDEs.

The paper is organized as follows. In Sect. 2 we present the model and the optimization problem. In Sect. 3 we transform the original problem into a BSPDVI and provide the verification problem and in Sect. 4 we verify the assumptions of the verification theorem. In Sect. 5 we study properties of the optimal boundary and in Sect. 6 we conclude. The Appendix gives existing results on BSPDVI and derive technical lemmas, which are necessary to provide a characterization of the optimal retirement boundary.

## 2 The Model

We consider an agent whose objective is to maximize the following time-separable von Neumann–Morgenstern utility function by choosing her consumption/leisure, the portfolio of investments and the retirement time:

$$\mathbb{E}\left[\int_{0}^{\tau} e^{-\beta t} U_{1}(c_{t}, l_{t}) dt + \int_{\tau}^{T} e^{-\beta t} U_{1}(c_{t}, L_{t}) dt + e^{-\beta T} U_{2}(Y_{T})\right], \quad (2.1)$$

where  $\tau \in [0, T]$  is the time of the agent's retirement, and the constant T denotes the fixed mandatory retirement date, and  $\beta$  is her subjective constant discount rate, and  $Y_T$  is her wealth at time T. There exists a single consumption good and  $c_t$  denotes the agent's rate of consumption at time t. The agent derives utility from enjoying leisure, and  $l_t$  denotes the rate of her taking leisure at t satisfying the condition  $0 \le l_t \le \overline{L}_t < L_t$ , where  $\overline{L}$  and L are two positive stochastic processes satisfying Assumption 2 in Sect. 2.1. The model is that of a flexible labor supply similar to Bodie et al. [3], i.e., the agent's labor supply is not fixed, but can change over time according to her choice: L denotes the total leisure endowed to the agent, and L - l is the rate at which she supplies labor, i.e., by sacrificing part of the endowed leisure, the agent supplies her labor. There is a minimum work requirement before retirement, i.e., she is required to work at a rate greater than or equal to  $L - \overline{L}$ .

The function  $U_2(\cdot)$  is the utility function of wealth after retirement. The agent has an option to retire earlier than the mandatory retirement date T, and the retirement decision is irreversible, i.e., the agent does not have an option to go back to work once she chooses to retire. The agent receives income at a rate equal to  $w_t > 0$  for labor she supplies, and thus the rate of her labor income is equal to  $w_t (L_t - l_t)$  at time t. After retirement, the agent enjoys her full endowed leisure but does not earn labor income. Thus, the choice of the optimal retirement time  $\tau$  hinges on the trade-off between the pecuniary benefit of labor income and the marginal utility of enjoying leisure fully after retirement. Choi et al. [5] have considered a similar problem in a Markovian model with an infinite horizon.

#### Applied Mathematics & Optimization

For simplicity of our analysis we assume that  $U_1(\cdot, \cdot)$  is the following constant elasticity of substitution (CES) utility function as in Choi et al. [5]:

$$U_{1}(c,l) = \begin{cases} \frac{\left[\alpha c^{\rho} + (1-\alpha)l^{\rho}\right]^{\frac{1-\gamma}{\rho}}}{1-\gamma}, & 0 < \gamma \neq 1, 0 < \alpha < 1, 0 \neq \rho < 1; \\ \frac{1}{\rho} \log\left[\alpha c^{\rho} + (1-\alpha)l^{\rho}\right], & \gamma = 1, 0 < \alpha < 1, 0 \neq \rho < 1; \\ \frac{\left[c^{\alpha}l^{1-\alpha}\right]^{1-\gamma}}{1-\gamma} & 0 < \gamma \neq 1, 0 < \alpha < 1, \rho = 0; \\ \log\left(c^{\alpha}l^{1-\alpha}\right) & \gamma = 1, 0 < \alpha < 1, \rho = 0, \end{cases}$$

where  $\gamma$ ,  $\alpha$  and  $\rho$  are constants. Here  $\gamma$  is the coefficient of relative risk aversion or the reciprocal of the elasticity of intertemporal substitution,  $1/(1-\rho)$  is the elasticity of substitution between consumption and leisure,  $\alpha$  is the share of contribution of consumption to utility.

We also make the following assumption about  $U_2(\cdot)$ .

**Assumption 1** The utility function  $U_2(y) \in C^2(0, +\infty)$  taking values in  $\mathbb{R}^+ \triangleq (0, +\infty)$ , is strictly increasing and strictly concave with respect to y and satisfies that there exist two positive constants C and k such that<sup>2</sup>

$$\lim_{y \to 0^+} U_2'(y) = +\infty, \quad \lim_{y \to +\infty} U_2'(y) = 0, \quad \limsup_{y \to +\infty} U_2'(y) \ y^k \le C.$$

The first two conditions in Assumption 1 are called *Inada conditions*, which are commonly employed in models of economic growth and consumption/savings (see e.g., Uzawa [29], Inada [16] etc.). The last inequality in Assumption 1 is a technical assumption, which is satisfied by many commonly used utility functions. For example, it is satisfied by the following constant relative risk aversion (CRRA) utility function and constant absolute risk aversion (CARA) utility function, i.e.,

$$U_2(y) = \frac{y^{1-\gamma}}{1-\gamma} \ (0 < \gamma \neq 1) \ \text{ or } \ U_2(y) = \log y \ \text{ or } \ U_2(y) = 1 - e^{-\alpha y} \ (\alpha > 0).$$

Remark 1 From Assumption 1, we can deduce that

$$0 \le \mathcal{J}_{U_2}(x) \le C(1 + x^{-1/k}), \quad \forall \ x \in (0, +\infty),$$
(2.2)

where  $\mathcal{J}_{U_2}(\cdot)$  is the inverse function of  $U'_2(\cdot)$ , and C is a positive constant.

Yang and Koo [30] and Koo et al. [21] considered a felicity function of the form U(c) - l where l is disutility of labor, and their work is different from the model in this paper since they do not allow flexible labor supply.

 $<sup>^2</sup>$  In order to save notation we will use the same notation *C* (and *K*) to denote different large enough positive constants. Their meanings, however, will be clear from the context in which they appear.

#### 2.1 Financial Market

There are one riskless asset and  $(N_1 + N_2)$  risky assets in the financial market of the economy. The price  $P_0$  of the riskless asset and the price  $P_i$  of the *i*-th risky asset are governed by the following stochastic differential equation (SDE)<sup>3</sup>

$$\begin{cases} P_{0,t} = P_0 + \int_0^t r_u P_{0,u} du; \\ P_{i,t} = P_i + \int_0^t \mu_{i,u} P_{i,u} du + \int_0^t \sum_{j=1}^{N_1} \sigma_{ij,u}^1 P_{i,u} dW_{j,u} + \int_0^t \sum_{j=1}^{N_2} \sigma_{ij,u}^2 P_{i,u} dB_{j,u}. \end{cases}$$

where  $i = 1, 2, \dots, N_1 + N_2$ . The sources of risk are described by two independent Brownian motions:  $N_1$ -dimensional standard Brownian motion,  $W = (W_1, \dots, W_{N_1})^{\top}$  and  $N_2$ -dimensional standard Brownian motion,  $B = (B_1, \dots, B_{N_2})^{\top}$ , where  $A^{\top}$  denotes the transpose of matrix A. The prices of the risky assets are described by the  $(N_1 + N_2)$ -dimensional stochastic process  $P = (P_1, \dots, P_{N_1+N_2})^{\top}$ . In the SDE for  $P_0$ , r represents the risk-free interest rate. And in the SDE for P,  $\mu = (\mu_1, \dots, \mu_{N_1+N_2})^{\top}$ ,  $\Sigma = (\Sigma^1, \Sigma^2)$  represent the means and the sensitivity of the returns on the risky assets to risk sources, respectively, where  $\Sigma^1 = (\sigma_{ij}^1)_{(N_1+N_2)\times N_1}$ ,  $\Sigma^2 = (\sigma_{ij}^2)_{(N_1+N_2)\times N_2}$ . We assume that  $\Sigma$  is strongly non-degenerate, i.e., there exists a positive constant  $\kappa$  such that  $\xi^{\top} \Sigma \xi \ge \kappa |\xi|^2$  for any  $\xi \in \mathbb{R}^{N_1+N_2}$ . The assumption implies, in particular, that the financial market is complete.

The Brownian motion  $(B^{\top}, W^{\top})^{\top}$  is defined on a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F} \triangleq (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$ . Moreover, let us denote by  $\mathbb{F}^W \triangleq \{\mathcal{F}_t^W\}_{t=0}^T$  and  $\mathbb{F}^B \triangleq \{\mathcal{F}_t^B\}_{t=0}^T$ , respectively, the natural filtrations generated by W and B containing all  $\mathbb{P}$ -null sets in  $\mathcal{F}$ . Without loss of generality, we assume that  $\mathbb{F} \triangleq \mathbb{F}^W \vee \mathbb{F}^B$ . We denote by  $\mathcal{P}$  and  $\mathcal{P}^B$  the  $\sigma$ - algebras of predictable sets in  $\Omega \times [0, T]$  associated with  $\mathbb{F}$  and  $\mathbb{F}^B$ , respectively.

Let us define the market price of risk  $\theta$  by

$$\theta \triangleq (\Sigma)^{-1} \Big( \mu - r \mathbf{1}_{N_1 + N_2} \Big), \quad i = 1, 2,$$

where  $1_{N_1+N_2}$  is the  $(N_1 + N_2)$ -dimensional column vector of 1's. Let us write  $\theta = ((\theta^1)^{\top}, (\theta^2)^{\top})^{\top}$ , where  $\theta^1$  and  $\theta^2$  are the first  $N_1$ -subvector and the last  $N_2$ -subvector of  $\theta$ , respectively. Then, we can define the state price density process H by

$$H_t \triangleq \exp\left\{-\int_0^t r_u \, du \, - \, \frac{1}{2} \int_0^t |\theta_u|^2 \, du - \int_0^t (\theta_u^1)^\top \, dW_u - \int_0^t (\theta_u^2)^\top \, dB_u\right\}. \tag{2.3}$$

<sup>&</sup>lt;sup>3</sup> In this model, we suppose that the risky assets are driven by two independent Brownian motions W and B, and the coefficient r,  $\mu$ ,  $\Sigma$  are only  $\mathcal{P}^B$ -measurable ( $\mathcal{P}^B$  is the  $\sigma$ -algebra of predictable sets generated by B). We make this assumption to ensure that the differential operator in the associated BSPDE and BSPDVI is non-degenerate, and the BSPDE and BSPDVI have strong solutions. In fact, our model includes such commonly used models as some stochastic interest rate models and stochastic volatility models.

#### Applied Mathematics & Optimization

Throughout this paper, we suppose that the stochastic processes  $\overline{L}$ , L and the coefficients r,  $\theta^1$ ,  $\theta^2$  satisfy the following Assumptions 2 and 3:

**Assumption 2** (*Measurability and Boundedness*) The stochastic processes  $\overline{L}$ , L and the coefficient processes r,  $\theta^1$ ,  $\theta^2$ , the volatility process  $\Sigma$ , and the income rate process w are  $\mathcal{P}^B$ -measurable with values in  $\mathbb{R}^+$ ,  $\mathbb{R}^+$ ,  $\mathbb{R}^1$ ,  $\mathbb{R}^{N_1}$ ,  $\mathbb{R}^{N_2}$ ,

 $\mathbb{R}^{(N_1+N_2)\times(N_1+N_2)}$ , and  $\mathbb{R}^+$ , respectively. Moreover, there exists a positive constant *C* such that

$$L + \frac{1}{\overline{L}} + \frac{1}{L - \overline{L}} + |r| + |\theta^{1}| + |\theta^{2}| + |\Sigma| + w + \frac{1}{w} \le C \text{ a.e. in } \Omega \times [0, T].$$

According to Assumption 2 the risk-free rate r and the market price of risk  $\theta$  are  $\mathcal{P}^B$ -measurable, and thus, change stochastically according to the movement of Brownian motion B. Thus, the Brownian motion B can be regarded as sources of risk which cause changes in the investment opportunity. Notice that we do not impose Markovian restrictions on the changes.

**Assumption 3** (*Nondegeneracy*) There exists a positive constant  $\kappa$  such that  $|\theta^1| \ge \kappa$  a.e. in  $\Omega \times [0, T]$ .

From the expression of *H* in (2.3), we know that  $H \in S^p$  for any  $p \ge 1$  under Assumption 2 (see Subsection 2.3 for the definition of  $S^p$ ).

#### 2.2 Optimization Problem

Let us denote the agent's monetary amounts of investment in the risky assets by  $\pi = (\pi_1, \dots, \pi_{N_1+N_2})^{\mathsf{T}}$ . Suppose that the current time is  $t \in [0, T]$  and the agent has not yet retired. Let  $\mathcal{U}_{t,T}$  be the class of all  $\mathbb{F}$ -stopping times which take values in [t, T]. The agent's wealth process Y is governed by

$$Y_{s}^{t,y;\tau,c,l,\pi} = y + \int_{t}^{s} \left[ \pi_{u}^{\top} \left( \mu_{u} - r_{u} \mathbf{1}_{N_{1}+N_{2}} \right) + r_{u} Y_{u}^{t,y;\tau,c,l,\pi} - c_{u} + w_{u} (L_{u} - l_{u}) I_{\{u \leq \tau\}} \right] du + \int_{t}^{s} \pi_{u}^{\top} \left[ \Sigma_{u}^{1} dW_{u} + \Sigma_{u}^{2} dB_{u} \right], \quad (2.4)$$

where y is a  $\mathcal{F}_t^B$ -measurable<sup>4</sup> random variable, and  $I_A$  is the characteristic function of a set  $A \subset \Omega$ .

$$\mathcal{Y}_{s} = -\int_{s}^{T} \left[ (Z_{u}^{\mathcal{Y}})^{\top} \theta_{u} + r_{u} \mathcal{Y}_{u} - L_{u} w_{u} \right] du - \int_{s}^{T} (Z_{u}^{\mathcal{Y}})^{\top} (dW_{u}^{\top}, dB_{u}^{\top})^{\top}.$$
(2.5)

Note that  $L, \theta, r, w$  are  $\mathcal{P}^B$ -measurable. Hence,  $\mathcal{Y}, Z^{\mathcal{Y}}$  are  $\mathcal{P}^B$ -measurable, and  $Z_i^{\mathcal{Y}} = 0$  for  $i = 1, \dots, N_1$ .

<sup>&</sup>lt;sup>4</sup> The negative of the present value  $-\mathcal{Y}_t$  is the initial level of wealth such that  $Y_T^{t,-\mathcal{Y}_t;T,0,0,\pi^{\mathcal{Y}}} = 0$  with an appropriate investment strategy  $\pi^{\mathcal{Y}}$ . So, if we let  $\mathcal{Y} = -Y^{t,-\mathcal{Y}_t;T,0,0,\pi^{\mathcal{Y}}}$ ,  $Z^{\mathcal{Y}} = -\Sigma^{\top}\pi^{\mathcal{Y}}$ , then  $(\mathcal{Y}, Z^{\mathcal{Y}})$  is the unique solution of the following backward stochastic differential equation (BSDE)

Since the number of risky assets is equal to the number of sources of risk, the financial market is complete and the present value  $\mathcal{Y}_t$  of cash flows can be calculated by using the state price density (see e.g., Karatzas et al. [18], Cox and Huang [6]). We define the present value  $\mathcal{Y}_t$  of labor income at *t* under the assumption that the agent provides a maximum possible supply of labor, *L*, and does not retire until *T* as in the following,

$$\mathcal{Y}_t = \mathbb{E}\left[\left|\int_t^T H_s^t L_s w_s ds\right| \mathcal{F}_t\right], \qquad H_s^t \triangleq \frac{H_s}{H_t} > 0 \quad \text{for } s \ge t.$$
(2.6)

It is clear that  $H_s^t$  satisfies the following SDE,

$$H_{s}^{t} = 1 - \int_{t}^{s} r_{u} H_{u}^{t} du - \int_{t}^{s} H_{u}^{t} (\theta_{u}^{1})^{\top} dW_{u} - \int_{t}^{s} H_{u}^{t} (\theta_{u}^{2})^{\top} dB_{u}.$$
 (2.7)

From Assumption 2, we see that  $r, \theta, w$  are bounded. Recalling  $L, w \ge 0$  and (2.6) and (2.7), by the theory of SDEs (see e.g., Mao [23]), we derive the following estimation:

$$0 \leq \mathcal{Y}_t \leq \|Lw\|_{\infty} \mathbb{E}\left[\int_t^T H_s^t ds \Big| \mathcal{F}_t\right] \leq C, \qquad (2.8)$$

where C is a constant independent of t, and we have used the notation

$$||Lw||_{\infty} \triangleq \operatorname{ess.sup}\{|L_u w_u| : (\omega, u) \in \Omega \times [0, T]\},$$

which will be used throughout this paper.

A policy,  $(\tau, c, l, \pi)$ , is admissible if  $\tau \in U_{t,T}$ , the set of all  $\mathbb{F}$ -stopping times taking values in [t, T], and  $c, l, \pi$  are  $\mathcal{P}$ -measurable, and it satisfies c > 0;  $0 < l_s \leq \overline{L_s}$  for any  $t \leq s \leq \tau, l_s = L_s$  for any  $\tau < s \leq T$ , and

$$\int_t^T (c_s + |\pi_s|^2) ds < \infty \text{ subject to } Y_s^{t,y;\tau,c,l,\pi} > -\mathcal{Y}_s I_{\{s<\tau\}} \text{ a.s. in } \Omega.$$

We will denote the set of all admissible policies by A(t, y).

The agent's problem at time t is to maximize

$$J(t, y; \tau, c, l, \pi) \triangleq \mathbb{E} \bigg[ \int_t^\tau e^{-\beta(u-t)} U_1(c_u, l_u) du + \int_\tau^T e^{-\beta(u-t)} U_1(c_u, L_u) du \\ + e^{-\beta(T-t)} U_2(Y_T^{t,y;\tau,c,l,\pi}) \bigg| \mathcal{F}_t \bigg],$$

where  $y > -\mathcal{Y}_t$ . That is, the agent would like to find an optimal strategy  $(\tau^*, c^*, l^*, \pi^*) \in \mathcal{A}^1(t, y)$  such that,

$$J(t, y; \tau^*, c^*, l^*, \pi^*) = V_t(y) \triangleq \operatorname{ess.sup} \left\{ J(t, y; \tau, c, l, \pi) : (\tau, c, l, \pi) \in \mathcal{A}^1(t, y) \right\},\$$

where

$$\begin{split} \mathcal{A}^{1}(t, y) &\triangleq \left\{ \left(\tau, c, l, \pi\right) \in \mathcal{A}(t, y) : \mathbb{E}\left[\int_{t}^{\tau} e^{-\beta(u-t)} U_{1}^{-}(c_{u}, l_{u}) du \right. \\ \left. + \left. \int_{\tau}^{T} e^{-\beta(u-t)} U_{1}^{-}(c_{u}, L_{u}) du + e^{-\beta(T-t)} U_{2}^{-}(Y_{T}^{t, y; \tau, c, l, \pi}) \right| \mathcal{F}_{t} \right] < + \infty \right\} \end{split}$$

with  $U^{-} \triangleq \max\{0, -U\}$ . And V is called *the value* of the optimization problem.

#### 2.3 Notation for Spaces of Stochastic Processes and Functions

In order to facilitate exposition of the paper, we introduce notation for spaces of stochastic processes and function spaces. We refer to [7,8] for more details of the function spaces.

For an integer  $m \in \mathbb{N}$ ,  $p \in [1, +\infty)$ ,  $\lambda \in [0, +\infty)$ , a smooth domain *D* in  $\mathbb{R}$ , we introduce the following spaces:

•  $C^m(D)$ , the set of all functions  $\eta : D \to E$  such that  $\eta$  and  $\eta', \eta'', \dots, \eta^{(m)}$  are continuous, where E is  $\mathbb{R}$  or  $\mathbb{R}^{N_2}$ ;

- $C_0^m(D)$ , the set of all functions in  $C^m(D)$  with compact supports in D;
- $H_{\lambda}^{m, p}(D)$ , the completion of  $C^{m}(D)$  under the norm

$$|\eta|_{m, p; \lambda} \triangleq \left(\int_{D} |\eta|^{p} e^{-\lambda|x|} dx + \sum_{i=1}^{m} \int_{\mathbb{R}} |\eta^{(i)}|^{p} e^{-\lambda|x|} dx\right)^{\frac{1}{p}};$$

•  $\mathbb{L}^{m,p}_{\lambda}(D)$ , the set of all  $H^{m,p}_{\lambda}(D)$ -valued and  $\mathcal{F}^B_T$ -measurable random variables such that

$$\mathbb{E}\left[\left\|\varphi\right\|_{m,p;\,\lambda}^{p}\right] < \infty;$$

•  $\mathcal{L}^{p}$ , the set of all  $\mathcal{P}$ -predictable stochastic processes taking values in  $\mathbb{R}$  with norm

$$\|X\|_{p} \triangleq \left\{ \mathbb{E}\left[\int_{0}^{T} |X_{t}|^{p} dt\right] \right\}^{\frac{1}{p}};$$

•  $S^p$ , the set of all path continuous processes in  $\mathcal{L}^p$  with norm

$$|||X|||_{p} \triangleq \left\{ \mathbb{E}\left[\sup_{t \in [0, T]} |X_{t}|^{p}\right] \right\}^{\frac{1}{p}};$$

•  $L^p(\mathcal{F}_T)$ , the space of  $\mathcal{F}_T$ -measurable random variables with norm  $\{\mathbb{E}[|X_T|^p]\}^{1/p}$  for random variable  $X_T$ .

•  $\mathbb{H}^{m, p}_{\lambda}(D)$ , the set of all  $\mathcal{P}^{B}$ -predictable stochastic processes with values in  $H^{m, p}_{\lambda}(D)$  with norm

$$\|V\|_{m, p; \lambda} \triangleq \left\{ \mathbb{E}\left[\int_0^T |V_t|_{m, p; \lambda}^p dt\right] \right\}^{\frac{1}{p}};$$

•  $\mathbb{S}_{\lambda}^{m, p}(D)$ , the set of all path continuous stochastic processes in  $\mathbb{H}_{\lambda}^{m, p}(D)$  equipped with norm

$$|||V|||_{m, p; \lambda} \triangleq \left\{ \mathbb{E} \left[ \sup_{t \in [0, T]} |V_t|_{m, p; \lambda}^p \right] \right\}^{\frac{1}{p}};$$

The notation for spaces  $H_0^{m, p}(D)$ ,  $\mathbb{L}_0^{m, p}(D)$ ,  $\mathbb{H}_0^{m, p}(D)$  and  $\mathbb{S}_0^{m, p}(D)$  will be abbreviated as  $H^{m, p}(D)$ ,  $\mathbb{L}^{m, p}(D)$ ,  $\mathbb{H}^{m, p}(D)$  and  $\mathbb{S}^{m, p}(D)$  if there is no confusion. Moreover, we will omit (D) if  $D = \mathbb{R}$ .

## 3 Transformation of the Original Problem into a BSPDVI and Verification Theorem

In this section we will recast the original optimal stochastic control problem into a BSPDVI by making three successive transformations.<sup>5</sup> Next we will present the verification theorem (Theorem 1), which provides a connection to all transformations. The verification theorem states that the value and the optimal strategy of the original optimization problem can be described by the unique solution to BSPDVI (3.16).

#### 3.1 Static Budget Constraints and Convex Dual Functions

From the SDE (2.7) for the state-price-density H and SDEs (2.4), (2.5), we deduce the static budget constraint as follows:

$$\mathbb{E}\left[\left|H_{s}^{t}(Y_{s}^{t,y;\tau,c,l,\pi}+\mathcal{Y}_{s})+\int_{t}^{s}H_{u}^{t}(c_{u}+w_{u}l_{u})du\right|\mathcal{F}_{t}\right] \leq y+\mathcal{Y}_{t} \text{ if } 0 \leq t \leq s \leq \tau;$$
(3.1)

$$\mathbb{E}\left[\left|H_{s}^{t}Y_{s}^{t,y;\tau,c,l,\pi}\right|+\int_{t}^{s}H_{u}^{t}c_{u}du\right|\mathcal{F}_{t}\right] \leq y \text{ if } \tau \leq t \leq s \leq T.$$
(3.2)

As a preparation for the transformations we first introduce the convex dual functions of  $U_1(\cdot, \cdot)$ ,  $U_2(\cdot)$  such that

$$\widetilde{U}_{1,t}(x) = \sup_{c \ge 0} \{ U_1(c, L_t) - xc \} = U_1(\mathcal{J}_{U_1}(x; L_t), L_t) - x\mathcal{J}_{U_1}(x; L_t); \quad (3.3)$$

<sup>&</sup>lt;sup>5</sup> See Yang and Koo [30] for similar transformations in a simpler context in a Markovian model with constant coefficients.

$$\begin{aligned} \widetilde{U}_{2}(x) &= \sup_{y \ge 0} \{ U_{2}(y) - xy \} = U_{2}(\mathcal{J}_{U_{2}}(x)) - x\mathcal{J}_{U_{2}}(x); \\ \widehat{U}_{1,t}(x) &= \sup_{c \ge 0, 0 \le l \le \overline{L}_{t}} \{ U_{1}(c,l) - x(c+w_{t}l) \} \\ &= A_{t}(x)I_{\{x \ge \overline{x}_{t}\}} \\ &+ \left[ U_{1}(\mathcal{J}_{U_{1}}(x;\overline{L}_{t}),\overline{L}_{t}) - x(\mathcal{J}_{U_{1}}(x;\overline{L}_{t}) + \overline{L}_{t}w_{t}) \right] I_{\{0 < x < \overline{x}_{t}\}}, \end{aligned}$$
(3.4)

where  $\mathcal{J}_{U_1}(\cdot; l)$  is the inverse function of  $\partial_c U_1(\cdot, l)$ ,  $\mathcal{J}_{U_2}(\cdot)$  is the inverse function of  $U'_2(\cdot)$ , and

$$A_{t}(x) \triangleq \begin{cases} \frac{\gamma}{1-\gamma} \alpha^{\frac{1-\gamma}{\rho\gamma}} x^{\frac{\gamma-1}{\gamma}} \\ \times \left[ 1 + w_{t} \left( \frac{\alpha w_{t}}{1-\alpha} \right)^{\frac{1}{\rho-1}} \right]^{\frac{(1-\rho)(1-\gamma)}{\rho\gamma}}, \ 0 < \gamma \neq 1, 0 < \alpha < 1, 0 \neq \rho < 1; \\ \log \left\{ \frac{\alpha^{\frac{1}{\rho}}}{x} \left[ 1 + w_{t} \right]^{\frac{1-\rho}{\gamma}} \\ \times \left( \frac{\alpha w_{t}}{1-\alpha} \right)^{\frac{1}{\rho-1}} \right]^{\frac{1-\rho}{\rho}} \right\} - 1, \qquad \gamma = 1, 0 < \alpha < 1, 0 \neq \rho < 1; \\ \left[ \frac{\gamma}{1-\gamma} \alpha^{\frac{1-\gamma}{\gamma}} \left( \frac{\alpha w_{t}}{1-\alpha} \right)^{\frac{-(1-\gamma)(1-\alpha)}{\gamma}} x^{\frac{\gamma-1}{\gamma}}, \ 0 < \gamma \neq 1, 0 < \alpha < 1, \rho = 0; \\ \log \left[ \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{w_{t}^{1-\alpha}x} \right] - 1, \qquad \gamma = 1, 0 < \alpha < 1, \rho = 0, \end{cases}$$

and

$$\overline{x}_{t} \triangleq \begin{cases} \alpha^{\frac{1-\gamma}{\rho}} \left(\frac{\alpha w_{t}}{1-\alpha}\right)^{\frac{\gamma}{\rho-1}} \overline{L}_{t}^{-\gamma} \\ \times \left[1 + w_{t} \left(\frac{\alpha w_{t}}{1-\alpha}\right)^{\frac{1}{\rho-1}}\right]^{\frac{1-\rho-\gamma}{\rho}}, \quad \gamma > 0, 0 < \alpha < 1, 0 \neq \rho < 1; \quad (3.6) \\ \alpha \left(\frac{\alpha w_{t}}{1-\alpha}\right)^{\alpha(1-\gamma)-1} \overline{L}_{t}^{-\gamma}, \quad \gamma > 0, 0 < \alpha < 1, \rho = 0. \end{cases}$$

Moreover, from the expression of  $U_1$ , Assumption 1, Assumption 2 and (2.2) we can derive the following properties of  $\widetilde{U}_1$ ,  $\widehat{U}_1$ ,  $\widetilde{U}_2$ .

**Lemma 1** 1.  $\widetilde{U}_1, \widehat{U}_1, \widetilde{U}_2$  are strictly decreasing and convex with respect to x. And  $\widetilde{U}_2 \in C^2(0, +\infty)$ , and  $\widetilde{U}_1, \widehat{U}_1 \in C(0, +\infty)$  a.e in  $\Omega \times [0, T]$ . 2. There exist positive constants C and K such that

$$\begin{aligned} |\widetilde{U}_{1,t}(x)| + |\partial_x \widetilde{U}_{1,t}(x)| + |\widetilde{U}_2(x)| + |\widetilde{U}_2'(x)| + |\widehat{U}_{1,t}(x)| \\ + |\partial_x \widehat{U}_{1,t}(x)| \le C \Big( x^K + x^{-K} \Big) \end{aligned}$$
(3.7)

for any x > 0 a.e. in  $\Omega \times [0, T]$ .

3. ess.sup{ $\partial_x \widetilde{U}_{1,t}(x) : (\omega, t) \in \Omega \times [0, T]$ },  $\widetilde{U}'_2(x)$ , ess.sup{ $\partial_x \widehat{U}_{1,t}(x) : (\omega, t) \in \Omega \times [0, T]$ }  $\rightarrow -\infty$  as  $x \rightarrow 0^+$ , and ess.inf{ $\partial_x \widetilde{U}_{1,t}(x) : (\omega, t) \in \Omega \times [0, T]$ },  $\widetilde{U}'_2(x)$ , ess.inf{ $\partial_x \widehat{U}_{1,t}(x) : (\omega, t) \in \Omega \times [0, T]$ }  $\rightarrow 0^-$  as  $x \rightarrow +\infty$ .

#### 3.2 Optimization Problem after Retirement

In this subsection we will consider the agent's optimization problem after retirement. After retirement, the agent does not face any choice of a stopping time. Thus, the control does not involve stopping time  $\tau$ . Then, the admissible set is  $\mathcal{A}_t^1(t, y) \triangleq \{(c, l, \pi) : (t, c, l, \pi) \in \mathcal{A}^1(t, y)\}$ , where the admissible set is dependent on the initial time *t* and the initial wealth *y*. Let us denote the agent's value after retirement by  $\underline{V}$ , i.e.,

$$\underline{V}_{t}(y) = \underset{(c,l,\pi)\in\mathcal{A}_{t}^{1}(t,y)}{\text{ess.sup}} \mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} U_{1}(c_{s},L_{s}) ds + e^{-\beta(T-t)} U_{2}(Y_{T}^{t,y;t,c,l,\pi}) \middle| \mathcal{F}_{t}\right].$$

Now (3.2) implies that for any  $t \in [0, T]$ , x > 0, y > 0,

$$\underline{V}_{t}(y) - xy \leq \underset{(c,l,\pi)\in\mathcal{A}_{t}^{1}(t,y)}{\operatorname{ess.sup}} \mathbb{E} \left[ \int_{t}^{T} \left[ e^{-\beta(s-t)} U_{1}(c_{s}, L_{s}) - x H_{s}^{t} c_{s} \right] ds \\
+ \left[ e^{-\beta(T-t)} U_{2} \left( Y_{T}^{t,y;t,c,l,\pi} \right) - x H_{T}^{t} Y_{T}^{t,y;t,c,l,\pi} \right] \left| \mathcal{F}_{t} \right] \\
\leq \mathbb{E} \left[ \int_{t}^{T} e^{-\beta(s-t)} \widetilde{U}_{1,s}(X_{s}) ds + e^{-\beta(T-t)} \widetilde{U}_{2}(X_{T}) \left| \mathcal{F}_{t} \right] \\
\leq \widetilde{V}_{t}(x)$$
(3.8)

with  $X_s = x e^{\beta(s-t)} H_s^t \in S^p$  for any  $p \ge 1$ .

**Remark 2** The Lagrange multiplier, x, and the dual variable at time s,  $X_s$ , represent the agent's marginal utility of wealth at time t and at time  $s \in [t, T]$ , respectively.  $\underline{\tilde{V}}_t(x)$  is called the *dual value* of the agent's optimization problem after retirement. From the verification theorem (Theorem 1 below), we know that under suitable conditions, for any y > 0, there exists a unique  $\mathcal{F}_t$ -measurable random variable x > 0 such that the inequalities in the above hold as equalities, and  $\underline{\tilde{V}}$  is the convex dual function of the concave function  $\underline{V}$ , i.e.,<sup>6</sup>

$$\begin{split} \underline{\widetilde{V}}_t(x) &= \sup_{y>0} \Big[ \underline{V}_t(y) - xy \Big], \\ \underline{V}_t(y) &= \inf_{x>0} \Big[ \underline{\widetilde{V}}_t(x) + xy \Big], \\ \forall t \in [0, T], \ x > 0, \ y > 0 \end{split}$$

<sup>&</sup>lt;sup>6</sup> In fact, we will show that  $\underline{V}$  and  $\underline{\widetilde{V}}$  are continuous with respect to y and x a.e. in  $\Omega \times [0, T]$  in Theorem 1, respectively. So, we use sup, inf rather than ess.sup, ess.inf here.

Thus, it is possible to deduce properties of  $\underline{V}$  through those of  $\underline{\widetilde{V}}$ .

Itô's formula implies that X is governed by the following SDE

$$X_{s} = x + \int_{t}^{s} (\beta - r_{u}) X_{u} \, du - \int_{t}^{s} X_{u} (\theta_{u}^{1})^{\top} \, dW_{u} - \int_{t}^{s} X_{u} (\theta_{u}^{2})^{\top} \, dB_{u}.$$
 (3.9)

Since the risk-free rate r and the market price of risk  $\theta$  are  $\mathcal{P}^B$ -measurable by Assumption2, i.e., change stochastically, driven by the Brownian motion  $B, \underline{\tilde{V}}$  is expected to satisfy the following BSPDE<sup>7</sup>

$$\begin{cases} d\underline{\widetilde{V}}_{t} = -\left(\mathcal{L}\underline{\widetilde{V}}_{t} + \sum_{i=1}^{N_{2}} \mathcal{M}_{i}\underline{\widetilde{Z}}_{i,t} + \widetilde{U}_{1,t}\right) dt \\ + \sum_{i=1}^{N_{2}} \underline{\widetilde{Z}}_{i,t} dB_{i,t} \text{ in } \Omega \times [0, T] \times \mathbb{R}^{+}; \\ \underline{\widetilde{V}}_{T}(x) = U_{2}(x) \text{ for any } x \in \mathbb{R}^{+} \text{ a.s in } \Omega, \end{cases}$$

$$(3.10)$$

where we have used the following notation, which we will use throughout the paper,

$$\mathcal{L} \triangleq \frac{1}{2} |\theta|^2 x^2 \partial_{xx} + (\beta - r) x \partial_x - \beta, \quad \mathcal{M}_i \triangleq -\theta_i^2 x \partial_x, \ i = 1, \cdots, N_2.$$
(3.11)

## 3.3 Transformations

In this subsection we make successive transformations to change the original problem into a BSPDVI.

#### **Transformation 1.**

In the first step we apply the dynamic programming principle to transform the problem into an optimal consumption, investment/retirement problem where the utility function after retirement is given by  $\underline{V}$  the value of the agent after retirement, which has been discovered in the previous subsection.

From the original problem we can deduce that

$$V_{t}(y) \leq \underset{(\tau,c,l,\pi)\in\mathcal{A}^{1}(t,y)}{\text{ess.sup}} \mathbb{E}\left[\int_{t}^{\tau} e^{-\beta(s-t)} U_{1}(c_{s},l_{s}) ds + e^{-\beta(\tau-t)} \underline{V}_{\tau}\left(Y_{\tau}^{t,y;\tau,c,l,\pi}\right) \middle| \mathcal{F}_{t}\right]$$

$$(3.12)$$

subject to (2.4).

*Remark 3* In fact, from the verification theorem (Theorem 1) below, we know that the inequality can be replaced by equality.

<sup>&</sup>lt;sup>7</sup> By [27]  $\underline{\tilde{V}}$  and  $\hat{V}$  defined in (3.15) are the viscosity solutions of BSPDE (3.10) and BSPDVI (3.16), respectively. In this paper, however, we do not need this result. We only focus on the verification theorem, i.e., we will construct the value V and the optimal strategies by means of the strong solutions of BSPDE (3.10) and BSPDVI (3.16) via Theorem 1 in Sect. 3.5.

#### **Transformation 2.**

In the second step we transform the original problem, which involves both stochastic control and optimal stopping, into a standard optimal stopping problem which does not involve stochastic control. We use the martingale and dual method, following the idea in Karatzas and Shreve [19] and Karatzas and Wang [20].

For a Lagrange multiplier x > 0, we define  $\widehat{\underline{V}}$  as

$$\underline{\widehat{V}}_t(x) = \underline{\widetilde{V}}_t(x) - x\mathcal{Y}_t \text{ in } \Omega \times [0, T] \times \mathbb{R}^+.$$
(3.13)

**Remark 4** There is a difference in income before and after retirement, and the term  $-x\mathcal{Y}_t$  is necessary to adjust the dual value after retirement to take into account the difference as will be shown below in (3.14).

So (3.1), (3.5), (3.8), (3.12) and (3.13) imply that

$$V_{t}(y) - x (y + \mathcal{Y}_{t})$$

$$\leq \underset{(\tau,c,l,\pi)\in\mathcal{A}^{1}(t,y)}{\operatorname{ess.sup}} \mathbb{E} \bigg[ \int_{t}^{\tau} e^{-\beta(s-t)} U_{1}(c_{s}, l_{s}) \, ds + e^{-\beta(\tau-t)} \underline{V}_{\tau} \Big( Y_{\tau}^{t,y;\tau,c,l,\pi} \Big) \\ - x \bigg[ \int_{t}^{\tau} H_{s}^{t}(c_{s} + w_{s}l_{s}) \, ds + H_{\tau}^{t} \Big( Y_{\tau}^{t,y;\tau,c,l,\pi} + \mathcal{Y}_{\tau} \Big) \bigg] \Big| \mathcal{F}_{t} \bigg] \\ = \underset{(\tau,c,l,\pi)\in\mathcal{A}^{1}(t,y)}{\operatorname{ess.sup}} \mathbb{E} \bigg[ \int_{t}^{\tau} \bigg[ e^{-\beta(s-t)} U_{1}(c_{s}, l_{s}) - x H_{s}^{t}(c_{s} + w_{s}l_{s}) \bigg] ds \\ + \bigg[ e^{-\beta(\tau-t)} \underline{V}_{\tau} \Big( Y_{\tau}^{t,y;\tau,c,l,\pi} \Big) - x H_{\tau}^{t} \Big( Y_{\tau}^{t,y;\tau,c,l,\pi} + \mathcal{Y}_{\tau} \Big) \bigg] \Big| \mathcal{F}_{t} \bigg] \\ \leq \underset{\tau\in\mathcal{U}_{t,T}}{\operatorname{ess.sup}} \mathbb{E} \bigg[ \int_{t}^{\tau} e^{-\beta(s-t)} \widehat{U}_{1,s} (X_{s}) ds + e^{-\beta(\tau-t)} \Big( \underbrace{\widetilde{V}}_{\tau}(X_{\tau}) - X_{\tau} \mathcal{Y}_{\tau} \Big) \Big| \mathcal{F}_{t} \bigg] \\ = \underset{\tau\in\mathcal{U}_{t,T}}{\operatorname{ess.sup}} \mathbb{E} \bigg[ \int_{t}^{\tau} e^{-\beta(s-t)} \widehat{U}_{1,s} (X_{s}) ds + e^{-\beta(\tau-t)} \Big( \underbrace{\widetilde{V}}_{\tau}(X_{\tau}) \Big| \mathcal{F}_{t} \bigg] \\ \end{aligned}$$

$$(3.15)$$

where we recall  $X_s = x e^{\beta(s-t)} H_s^t$ .

**Remark 5** In fact, for any  $y > -\mathcal{Y}_t$ , we conjecture that the inequalities in the above hold as equalities for a unique  $\mathcal{F}_t$ -measurable random variable x > 0. And

$$\widehat{V}_t(x) = \sup_{y > -\mathcal{Y}_t} \left[ V_t(y) - x(y + \mathcal{Y}_t) \right], \quad V_t(y) = \inf_{x > 0} \left[ \widehat{V}_t(x) + x(y + \mathcal{Y}_t) \right]$$

for any  $t \in [0, T]$ , x > 0,  $y > -\mathcal{Y}_t$  a.s. in  $\Omega$ . If the conjecture is true, then we can derive properties of *V* through those of  $\widehat{V}$ . Moreover, we will show that the conjecture is true in the verification theorem (Theorem 1).

#### Applied Mathematics & Optimization

#### **Transformation 3.**

The optimization problem represented by the right-hand side of the last equality in (3.14) is a standard optimal stopping problem for  $t \in [0, T]$ . Thus, in this last step we can use the relationship between optimal stopping problems and BSPDVIs to transform the original problem into a BSPDVI (see e.g., [21,31]):

$$\begin{cases} d\widehat{V}_{t} = -\left(\mathcal{L}\widehat{V}_{t} + \sum_{i=1}^{N_{2}}\mathcal{M}_{i}\widehat{Z}_{i,t} + \widehat{U}_{1,t}\right) dt + \sum_{i=1}^{N_{2}}\widehat{Z}_{i,t} dB_{i,t} & \text{if } \widehat{V} > \underline{\widehat{V}}; \\ d\widehat{V}_{t} \leq -\left(\mathcal{L}\widehat{V}_{t} + \sum_{i=1}^{N_{2}}\mathcal{M}_{i}\widehat{Z}_{i,t} + \widehat{U}_{1,t}\right) dt + \sum_{i=1}^{N_{2}}\widehat{Z}_{i,t} dB_{i,t} & \text{if } \widehat{V} = \underline{\widehat{V}}; \\ \widehat{V}_{T}(x) = \underline{\widehat{V}}_{T}(x) & \text{for any } x \in \mathbb{R}^{+} \text{ a.s in } \Omega. \end{cases}$$

$$(3.16)$$

## 3.4 Strong Solution to BSPDE or BSPDVI

In this subsection, we introduce the definition of the strong solution to BSPDE or BSPDVI, the details can be found in [8,21,31].

**Definition 1** If the two-tuple  $(\tilde{V}, \tilde{Z}) \in \mathbb{H}^{2,2}(D) \times \mathbb{H}^{1,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ , and satisfies

$$\underline{\widetilde{V}}_{t} = \widetilde{U}_{2} + \int_{t}^{T} \left( \mathcal{L}\underline{\widetilde{V}}_{s} + \sum_{i=1}^{N_{2}} \mathcal{M}_{i}\underline{\widetilde{Z}}_{i,s} + \widetilde{U}_{1,s} \right) ds - \sum_{i=1}^{N_{2}} \int_{t}^{T} \widetilde{Z}_{i,s} dB_{i,s}, \quad (3.17)$$

a.e.  $x \in \mathbb{R}^+$  for all  $t \in [0, T]$  and a.s. in  $\Omega$ . Then  $(\underline{\widetilde{V}}, \underline{\widetilde{Z}})$  is called a strong solution of BSPDE (3.10).

**Definition 2** If the triplet  $(\widehat{V}, \widehat{Z}, \widehat{k}^+) \in \mathbb{H}^{2,2}(D) \times \mathbb{H}^{1,2}(D) \times \mathbb{H}^{0,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ , and satisfies

$$\begin{cases} \widehat{V}_{t} = \underline{V}_{T} + \int_{t}^{T} \left( \mathcal{L}\widehat{V}_{s} + \sum_{i=1}^{N_{2}} \mathcal{M}_{i}\widehat{Z}_{i,s} + \widehat{U}_{1,s} + \widehat{k}_{s}^{+} \right) ds \\ -\sum_{i=1}^{N_{2}} \int_{t}^{T} \widehat{Z}_{i,s} dB_{i,s}, \text{ a.e. } x \in \mathbb{R}^{+} \text{ for all } t \in [0, T], \text{ a.s. in } \Omega; \quad (3.18) \\ \widehat{V} \geq \underline{\widehat{V}}, \quad \widehat{k}^{+} \geq 0 \quad \text{ a.e. in } \Omega \times [0, T] \times \mathbb{R}^{+}; \\ \int_{0}^{T} (\widehat{V}_{t} - \underline{\widehat{V}}_{t}) \widehat{k}^{+} dt = 0 \quad \text{ a.e. in } \Omega \times \mathbb{R}^{+}. \end{cases}$$

Then  $(\widehat{V}, \widehat{Z}, \widehat{k}^+)$  is called a strong solution to BSPDVI (3.16).

*Remark 1* We have modified slightly the definition of the strong solution from that in [21,31], where the strong solutions  $(\widehat{V}, \widehat{Z}, \widehat{k}^+) \in \mathbb{H}^{2,2}_{\lambda}(\mathbb{R}) \times \mathbb{H}^{1,2}_{\lambda}(\mathbb{R}) \times \mathbb{H}^{0,2}_{\lambda}(\mathbb{R}).$ 

Springer

The modification is not essential as we can transform the domain from  $\mathbb{R}^+$  into  $\mathbb{R}$  via  $\tilde{x} = \log x$  in (4.1), and utilize a sequence of intervals to  $\{D_n\}_{n=1}^{\infty}$ , approximate to  $\mathbb{R}^+$ , where  $D_n = [1/n, n]$ .

Moreover, by the theory for BSPDEs or BSPDVIs (refer to Lemmas 4 and 6, or [8,28,31]), we see that  $\underline{\tilde{V}}, \hat{V} \in \mathbb{S}^{1,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ . So,  $\underline{\tilde{V}}_t(\omega, \cdot), \hat{V}_t(\omega, \cdot) \in H^{1,2}(D)$  for any  $t \in [0, T]$  a.s. in  $\Omega$ . Thus the Sobolev embedding theory implies that  $\underline{\tilde{V}}, \hat{V}$  are continuous with respect to x for any  $t \in [0, T]$  a.s. in  $\Omega$ . Repeating a similar argument, by the fact that  $\underline{\tilde{V}}, \hat{V} \in \mathbb{H}^{2,2}(D)$ , we can show that  $\partial_x \underline{\tilde{V}}_t(\omega, x), \partial_x \hat{V}_t(\omega, x)$  are continuous with respect to x a.s. in  $\Omega \times [0, T]$ .

#### 3.5 Verification Theorem

In this subsection we will state and prove the verification theorem, which provides the value V and the optimal strategy  $(\tau^*, c^*, l^*, \pi^*)$  by using the strong solutions to BSPDVI (3.16) and BSPDE (3.10).

**Theorem 1** Suppose that  $(\tilde{v}, \tilde{z})$  is the strong solution to BSPDE (3.10), and denote  $\hat{v} = \tilde{v} - x\mathcal{Y}$ . Assume that  $(\hat{v}, \hat{z}, \hat{k}^+)$  is the strong solution to BSPDVI (3.16), where the lower obstacle  $\hat{V}$  is replaced by  $\hat{v}$ . Suppose that  $\hat{v}, \hat{v}$  have the following properties:

- 1.  $\partial_{xx}\hat{v} > 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ .
- 2.  $\partial_x \widehat{v}_t(x) \to -\infty \text{ as } x \to 0^+, \ \partial_x \widehat{v}_t(x) \to 0 \text{ as } x \to +\infty \text{ a.e. in } \Omega \times [0, T].$
- 3. There exist positive constants C, K such that

$$|\widehat{v}_t(x)| + |\widehat{v}_t(x)| \le C(x^K + x^{-K}) \text{ a.e in } \Omega \times [0, T] \times \mathbb{R}^+.$$

Then, the value V takes the form of

$$V_t(y) = \inf_{x>0} \left[ \hat{v}_t(x) + x(y + \mathcal{Y}_t) \right] = \hat{v}_t(x_t^*(y)) + x_t^*(y)(y + \mathcal{Y}_t), \quad (3.19)$$

where  $x_t^*(y) = \mathcal{J}_{\widehat{v},t}(-y - \mathcal{Y}_t) > 0$  a.e. in  $\{(\omega, t, y) : y > -\mathcal{Y}_t, (\omega, t) \in \Omega \times [0, T]\}$ , and  $\mathcal{J}_{\widehat{v},t}(\cdot)$  is the inverse function of  $\partial_x \widehat{v}_t(\cdot)$ .

The value V and  $\partial_y V = x^*$  are continuous with respect to y a.e. in  $\Omega \times [0, T]$ , and  $\partial_{yy} V < 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ .

Moreover,  $x^*$  is strictly decreasing with respect to y a.e. in  $\Omega \times [0, T]$ , and has the asymptotic properties:  $x_t^*(y) \to +\infty$  as  $y \to -\mathcal{Y}_t$ , and  $x_t^*(y) \to 0^+$  as  $y \to +\infty$  a.e. in  $\Omega \times [0, T]$ .

The optimal retirement strategy can be described as

$$\tau^* = \inf\{s \in [t, T] : \widehat{v}_s(X_s^*) = \underline{\widehat{v}}_s(X_s^*)\}, \quad X_s^* \triangleq x_t^*(y)e^{\beta(s-t)}H_s^t$$

and the optimal leisure rate, the optimal consumption rate can be described as

$$l_{s}^{*} = \overline{L}_{s} \min \left\{ 1, \left( \frac{\overline{x}_{t}}{X_{s}^{*}} \right)^{\frac{1}{\gamma}} \right\} I_{\{t \le s \le \tau^{*}\}} + L_{s} I_{\{\tau^{*} < s \le T\}}, \quad c_{s}^{*} = \mathcal{J}_{U_{1}}(X_{s}^{*}; l_{s}^{*})$$

🖄 Springer

where  $\overline{x}_t$  is defined in (3.6). Moreover, the optimal investment strategy  $\pi^*$  is governed by the following BSDE

$$Y_{s}^{*} = Y_{T}^{*} - \int_{s}^{T} \left[ (\pi_{u}^{*})^{\top} \Sigma_{u} \theta_{u} + r_{u} Y_{u}^{*} - c_{u}^{*} + w_{u} (L_{u} - l_{u}^{*}) I_{\{u \leq \tau^{*}\}} \right] du$$
$$- \int_{s}^{T} (\pi_{u}^{*})^{T} \left[ \Sigma_{u}^{1} dW_{u} + \Sigma_{u}^{2} dB_{u} \right], \quad \forall \ s \in [t, T], \quad Y_{T}^{*} = \mathcal{J}_{U_{2}}(X_{T}^{*}). \quad (3.20)$$

**Proof** First, we show that  $x_t^*(y)$  is well-defined and  $x_t^*(y) > 0$ . We also show that  $x_t^*(y)$  has the monotonicity and asymptotic properties as in the conclusion of the theorem.

Since  $\partial_{xx}\hat{v} > 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$  and  $\hat{v}, \partial_x \hat{v}$  are continuous with respect to *x* a.e. in  $\Omega \times [0, T]$  (refer to Remark 1), we deduce that  $\partial_x \hat{v}$  is strictly increasing with respect to *x* a.e. in  $\Omega \times [0, T]$ . And assumption 2 in this theorem implies that  $\partial_x \hat{v}_t(\omega, \cdot) : (0, +\infty) \to (-\infty, 0)$  a.e. in  $\Omega \times [0, T]$ . Hence, we deduce that  $x_t^*(y) = \mathcal{J}_{\hat{v},t}(-y - \mathcal{Y}_t)$  exists, takes values on  $(0, +\infty)$ , and is continuous and strictly decreasing with respect to *y* a.e. in  $\Omega \times [0, T]$ . And the asymptotic properties of  $x_t^*(y)$ come from those of  $\partial_x \hat{v}_t(x)$ .

Second, we show that that  $l^*$ ,  $c^*$ ,  $Y_T^*$  and  $\tau^*$  are well-defined. In fact,  $U_1(\cdot, l)$ ,  $U_2$  are strictly concave, and thus,  $\mathcal{J}_{U_1}(\cdot; l)$ ,  $\mathcal{J}_{U_2}$  are well-defined, and hence,  $l^*$ ,  $c^*$  and  $Y_T^*$  are well-defined. From Remark 6, we deduce that  $\hat{v}_s(X_s^*) - \underline{\hat{v}}_s(X_s^*)$  is continuous with respect to *s*. Thus, we know that  $\tau^*$  is a  $\mathbb{F}$ -stopping time. Moreover, the terminal value condition of BSPDVI (3.16) implies that  $\tau^* \leq T$  and  $\tau^* \in \mathcal{U}_{l,T}$ .

We will show in Lemma 2 that  $\pi^*$  can be constructed from the solution of BSDE (3.20) and  $(\tau^*, c^*, l^*, \pi^*) \in \mathcal{A}^1(t, Y_t^*)$ . Based on the lemma we continue the proof of the theorem. The optimal investment strategy  $\pi^*$  comes from BSDE (3.20) rather than SDE (2.4), and thus it is necessary to prove  $Y^* = Y^{t,y;\tau^*,c^*,l^*,\pi^*}$ . For this purpose, we compare SDE (2.4) with BSDE (3.20). We find that

$$d Y_s^{t,y;\tau^*,c^*,l^*,\pi^*} = d Y_s^*, \ \forall \ t \le s \le T.$$

Hence, the uniqueness of the solution of the SDE implies that it is sufficient to prove that  $Y_t^* = y$  a.s. in  $\Omega$ .

Applying Itô's formula, and recalling (3.20) and (3.9), we have

$$\begin{aligned} x_t^*(y)Y_t^* &= e^{-\beta(T-t)}X_T^*Y_T^* + \int_t^T e^{-\beta(u-t)}X_u^* \Big[ c_u^* - w_u(L_u - l_u^*)I_{\{u \le \tau^*\}} \Big] du \\ &- \int_t^T e^{-\beta(u-t)}X_u^* \Big[ (\pi_u^*)^\top \Sigma_u - \theta_u^\top Y_u^* \Big] (dW_u^\top, dB_u^\top)^\top. \end{aligned}$$

Taking  $\mathcal{F}_t$ -conditional expectation in this equality, and combining this with the fact that  $X^*$ ,  $Y^* \in S^p$ ,  $\pi^* \in \mathcal{L}^p$  for any  $p \ge 1$  (refer to the proof of Lemma 2 below), we have

 $x_t^*(y)Y_t^*$ 

Author's personal copy

$$= \mathbb{E}\left[e^{-\beta(T-t)}X_{T}^{*}Y_{T}^{*} + \int_{t}^{T}e^{-\beta(u-t)}X_{u}^{*}\left[c_{u}^{*} - w_{u}(L_{u} - l_{u}^{*})I_{\{u \leq \tau^{*}\}}\right]du \left|\mathcal{F}_{t}\right]\right]$$

$$= \mathbb{E}\left[e^{-\beta(T-t)}\left[U_{2}(Y_{T}^{*}) - \widetilde{U}_{2}(X_{T}^{*})\right]$$

$$+ \int_{t}^{T}e^{-\beta(u-t)}\left[\left(U_{1}(c_{u}^{*}, l_{u}^{*}) - \widehat{U}_{1,u}(X_{u}^{*})\right)I_{\{u \leq \tau^{*}\}}\right]$$

$$- L_{u}X_{u}^{*}w_{u}I_{\{u \leq \tau^{*}\}} + \left(U_{1}(c_{u}^{*}, L_{u}) - \widetilde{U}_{1,u}(X_{u}^{*})\right)I_{\{\tau^{*} < u \leq T\}}\right]du \left|\mathcal{F}_{t}\right]$$

$$= J(t, Y_{t}^{*}; \tau^{*}, c^{*}, l^{*}, \pi^{*}) - \mathbb{E}\left[e^{-\beta(T-t)}\widetilde{\underline{v}}_{T}(X_{T}^{*}) + \mathcal{K}^{1} + \mathcal{K}^{2}\left|\mathcal{F}_{t}\right].$$
(3.21)

Here we have used the definitions of  $c^*$ ,  $l^*$ ,  $Y_T^*$  in the second equality, and used the terminal condition of BSPDE (3.10) in the third equality, and we have used the notation

$$\mathcal{K}^{1} \triangleq \int_{t}^{\tau^{*}} e^{-\beta(u-t)} \left[ \widehat{U}_{1,u}(X_{u}^{*}) + L_{u}X_{u}^{*}w_{u} \right] du, \ \mathcal{K}^{2} \triangleq \int_{\tau^{*}}^{T} e^{-\beta(u-t)} \widetilde{U}_{1,u}(X_{u}^{*}) du.$$

Since  $\hat{v}_{\cdot}(X_{\cdot}^*)$  and  $\hat{v}_{\cdot}(X_{\cdot}^*)$  are continuous stochastic processes by Remark 6 in Appendix, we have

$$\widehat{v}_T(X_T^*) = \underline{\widehat{v}}_T(X_T^*) = \underline{\widetilde{v}}_T(X_T^*) - x\mathcal{Y}_T, \quad \widehat{v}_{\tau^*}(X_{\tau^*}^*) = \underline{\widehat{v}}_{\tau^*}(X_{\tau^*}^*).$$

Let us denote

$$D_n = [1/n, n], n \in \mathbb{Z}^+, \quad \tau_n = \{s \ge \tau^*, |\partial_x \widetilde{\underline{\nu}}_u(X_u^*)| + |\widetilde{\underline{z}}_{i,u}(X_u^*)| \le n\} \wedge T.$$

Since  $(\underline{\widetilde{v}}, \underline{\widetilde{z}}) \in \mathbb{H}^{2, 2}(D_n) \times \mathbb{H}^{1, 2}(D_n)$  for any  $n \in \mathbb{Z}^+$ , we claim that  $\tau_n \to T$  a.s. in  $\Omega$ . Moreover, applying the generalized Itô-Kunita-Wentzell's formula to  $\underline{\widetilde{v}}_{\cdot}(X_{\cdot}^*)$  in Lemma 3, we have

$$\begin{split} & \mathbb{E}\bigg[e^{-\beta(\tau_n-t)}\underline{\widetilde{v}}_{\tau_n}(X_{\tau_n}^*) + \int_{\tau^*}^{\tau_n} e^{-\beta(u-t)} \,\widetilde{U}_{1,u}(X_u^*) \,du \,\bigg|\mathcal{F}_t\bigg] \\ &= \mathbb{E}\bigg[e^{-\beta(\tau^*-t)}\underline{\widetilde{v}}_{\tau^*}(X_{\tau^*}^*) + \int_{\tau^*}^{\tau_n} e^{-\beta(u-t)} \left[-\left(\mathcal{L}\underline{\widetilde{v}}_u(x) + \sum_{i=1}^{N_2} \mathcal{M}_i\underline{\widetilde{z}}_{i,u}(x) + \widetilde{U}_{1,u}(x)\right) \right. \\ & \left. + \mathcal{L}\underline{\widetilde{v}}_u(x) + \sum_{i=1}^{N_2} \mathcal{M}_i\underline{\widetilde{z}}_{i,u}(x)\right]_{x=X_u^*} + \mathcal{K}^3 + \int_{\tau^*}^{\tau_n} e^{-\beta(u-t)}\widetilde{U}_{1,u}(X_u^*) \,du \,\bigg|\mathcal{F}_t\bigg] \\ &= \mathbb{E}\bigg[e^{-\beta(\tau^*-t)}\underline{\widetilde{v}}_{\tau^*}(X_{\tau^*}^*) + \mathcal{K}^3\,\bigg|\mathcal{F}_t\bigg], \end{split}$$

where we have used BSPDE (3.10) in the first equality, and the notation

$$\mathcal{K}^{3} \triangleq \int_{\tau^{*}}^{\tau_{n}} e^{-\beta(u-t)} \left\{ -\partial_{x} \widetilde{\underline{\upsilon}}_{u}(X_{u}^{*}) X_{u}^{*} \left[ (\theta_{u}^{1})^{\top} dW_{u} + (\theta_{u}^{2})^{\top} dB_{u} \right] + (\widetilde{\underline{z}}_{i,u}(X_{u}^{*}))^{\top} dB_{u} \right\}.$$

Recalling the definition of  $\tau_n$ , and  $X^* \in S^p$  for any  $p \ge 1$ ,  $\theta$  is bounded by the constant *C*, we deduce that  $\mathbb{E}[\mathcal{K}^3|\mathcal{F}_t] = 0$ , and

$$\mathbb{E}\bigg[e^{-\beta(\tau_n-t)}\widetilde{\underline{v}}_{\tau_n}(X^*_{\tau_n})+\int_{\tau^*}^{\tau_n}e^{-\beta(u-t)}\widetilde{U}_{1,u}(X^*_u)du\bigg|\mathcal{F}_t\bigg]=\mathbb{E}\bigg[e^{-\beta(\tau^*-t)}\widetilde{\underline{v}}_{\tau^*}(X^*_{\tau^*})\bigg|\mathcal{F}_t\bigg].$$

From (3.7), (2.8) and Assumption 3 in this theorem, we deduce that

$$|\underline{\widetilde{\nu}}_{\tau_n}(X^*_{\tau_n})| + |\widetilde{U}_{1,u}(X^*_u)| \le 2C \sup_{u \in [t,T]} (|X^*_u|^K + |X^*_u|^{-K}),$$

where *C* and *K* are independent of *n*. Recalling the fact that  $X^*$ ,  $1/X^* \in S^p$  for any  $p \ge 1$  (refer to the proof of Lemma 2 below), by the domain convergence theorem, we conclude that

$$\mathbb{E}\left[\left.e^{-\beta(T-t)}\widetilde{\underline{v}}_{T}(X_{T}^{*})+\mathcal{K}^{2}\right|\mathcal{F}_{t}\right]=\mathbb{E}\left[\left.e^{-\beta(\tau^{*}-t)}\widetilde{\underline{v}}_{\tau^{*}}(X_{\tau^{*}}^{*})\right|\mathcal{F}_{t}\right].$$

Substituting the above equality into (3.21), we obtain

$$J(t, Y_t^*; \tau^*, c^*, l^*, \pi^*) - x_t^*(y)Y_t^* = \mathbb{E}\bigg[e^{-\beta(\tau^*-t)}\widetilde{\underline{v}}_{\tau^*}(X_{\tau^*}^*) + \mathcal{K}^1 \bigg| \mathcal{F}_t\bigg].$$

Moreover, applying Itô's formula to  $e^{-\beta(\cdot-t)}X^*\mathcal{Y}$  and recalling (2.5), we have

$$\begin{aligned} x_t^*(y)\mathcal{Y}_t &= e^{-\beta(\tau^*-t)}X_{\tau^*}^*\mathcal{Y}_{\tau^*} + \int_t^{\tau^*} e^{-\beta(u-t)}L_uX_u^*w_u\,du \\ &+ \int_t^{\tau^*} e^{-\beta(u-t)}X_u^* \Big[ \theta_u^\top \mathcal{Y}_u + (\pi_u^{\mathcal{Y}})^\top \Sigma_u \Big] (dW_u^\top, dB_u^\top)^\top. \end{aligned}$$

Hence, we have

$$J(t, Y_{t}^{*}; \tau^{*}, c^{*}, l^{*}, \pi^{*}) - x_{t}^{*}(y) \Big( Y_{t}^{*} + \mathcal{Y}_{t} \Big)$$
  
=  $\mathbb{E} \bigg[ e^{-\beta(\tau^{*}-t)} \widetilde{\underline{v}}_{\tau^{*}}(X_{\tau^{*}}^{*}) + \mathcal{K}^{1} - e^{-\beta(\tau^{*}-t)} X_{\tau^{*}}^{*} \mathcal{Y}_{\tau^{*}}$   
 $- \int_{t}^{\tau^{*}} e^{-\beta(u-t)} L_{u} X_{u}^{*} w_{u} du \bigg| \mathcal{F}_{t} \bigg]$   
=  $\mathbb{E} \bigg[ e^{-\beta(\tau^{*}-t)} \widehat{v}_{\tau^{*}}(X_{\tau^{*}}^{*}) + \int_{t}^{\tau^{*}} e^{-\beta(u-t)} \widehat{U}_{1,u}(X_{u}^{*}) du \bigg| \mathcal{F}_{t} \bigg]$ 

$$= \widehat{v}_t(x_t^*(y)),$$

where we have used  $\underline{\hat{v}} = \underline{\tilde{v}} - x\mathcal{Y}$  in the second equality. And in the third equality, we have used the method similar to that in the above and the following equality

$$\widehat{v}_t = \underline{v}_{\tau^*} + \int_t^{\tau^*} \left( \mathcal{L}\widehat{v}_s + \sum_{i=1}^{N_2} \mathcal{M}_i \widehat{z}_{i,s} + \widehat{U}_{1,s} \right) ds - \sum_{i=1}^{N_2} \int_t^{\tau^*} \widehat{z}_{i,s} dB_{i,s},$$

which follows from the fact that  $\hat{v}$  is the strong solution to BSPDVI (3.16) and the definition of  $\tau^*$ .

Until now, we have proved that

$$J(t, Y_t^*; \tau^*, c^*, l^*, \pi^*) = \widehat{v}_t(x_t^*(y)) + x_t^*(y) \Big( Y_t^* + \mathcal{Y}_t \Big)$$
  

$$\geq \inf_{x>0} \Big\{ \widehat{v}_t(x) + x \Big( Y_t^* + \mathcal{Y}_t \Big) \Big\}$$
  

$$= \widehat{v}_t(\widehat{x}^*) + \widehat{x}^* \Big( Y_t^* + \mathcal{Y}_t \Big), \qquad (3.22)$$

where  $\widehat{x}^* = \mathcal{J}_{\widehat{v},t}(-Y_t^* - \mathcal{Y}_t).$ 

On the other hand, recalling (3.13), (3.8) and (3.10), and applying Lemma 3, we have

$$\widehat{\underline{V}}_{t}(x) = \mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} \widetilde{U}_{1,s}(X_{s}) ds + e^{-\beta(T-t)} \widetilde{\underline{v}}_{T}(X_{T}) \middle| \mathcal{F}_{t} \right] - x \mathcal{Y}_{t} \\
= \mathbb{E}\left[\widetilde{\underline{v}}_{t}(x) + \mathcal{K}^{4} \middle| \mathcal{F}_{t} \right] - x \mathcal{Y}_{t} = \widehat{\underline{v}}_{t}(x),$$
(3.23)

where

$$\mathcal{K}^4 \triangleq \int_t^T e^{-\beta(s-t)} \left\{ -\partial_x \widetilde{\underline{v}}_s(X_s) X_s \left[ (\theta_s^1)^\top dW_s + (\theta_s^2)^\top dB_s \right] + (\widetilde{\underline{z}}_{i,s}(X_s))^\top dB_s \right\}.$$

Combining (3.14), (3.23) and BSPDVI (3.16), and applying Lemma 3, we deduce that for any  $\tilde{y} > 0$ , x > 0,  $(\tau, c, l, \pi) \in \mathcal{A}^1(t, \tilde{y})$ ,

.

$$J(t, \widetilde{y}; \tau, c, l, \pi) \leq V_t(\widetilde{y}) \leq \widehat{V_t}(x) + x(\widetilde{y} + \mathcal{Y}_t)$$
  
$$\leq \underset{\tau \in \mathcal{U}_{t,T}}{\text{ess.sup}} \mathbb{E} \bigg[ \int_t^{\tau} e^{-\beta(s-t)} \widehat{U}_{1,s}(X_s) ds + e^{-\beta(\tau-t)} \widehat{v_{\tau}}(X_{\tau}) \bigg| \mathcal{F}_t \bigg]$$
  
$$+ x (\widetilde{y} + \mathcal{Y}_t)$$
  
$$\leq \underset{\tau \in \mathcal{U}_{t,T}}{\text{ess.sup}} \mathbb{E} \bigg[ \widehat{v_t}(x) + \mathcal{K}^5 \bigg| \mathcal{F}_t \bigg] + x(\widetilde{y} + \mathcal{Y}_t) = \widehat{v_t}(x) + x(\widetilde{y} + \mathcal{Y}_t),$$

#### Applied Mathematics & Optimization

where

$$\mathcal{K}^5 \triangleq \int_t^\tau e^{-\beta(s-t)} \left\{ -\partial_x \widehat{v}_s(X_s) X_s \left[ (\theta_s^1)^\top dW_s + (\theta_s^2)^\top dB_s \right] + (\widehat{z}_{i,s}(X_s))^\top dB_s \right\}.$$

Setting  $\tilde{y} = Y_t^*$ ,  $x = \hat{x}^*$ ,  $(\tau, c, l, \pi) = (\tau^*, c^*, l^*, \pi^*)$  in the inequality, and recalling (3.22), we deduce that

$$J(t, Y_{t}^{*}; \tau^{*}, c^{*}, l^{*}, \pi^{*}) = \widehat{v}_{t}(x_{t}^{*}(y)) + x_{t}^{*}(y) \left(Y_{t}^{*} + \mathcal{Y}_{t}\right)$$
  
$$= \inf_{x>0} \left\{ \widehat{v}_{t}(x) + x \left(Y_{t}^{*} + \mathcal{Y}_{t}\right) \right\}$$
  
$$= \widehat{v}_{t}(\widehat{x}^{*}) + \widehat{x}^{*} \left[Y_{t}^{*} + \mathcal{Y}_{t}\right]$$
  
$$= \widehat{V}_{t}(\widehat{x}^{*}) + \widehat{x}^{*} \left(Y_{t}^{*} + \mathcal{Y}_{t}\right)$$
  
$$= V_{t}(Y_{t}^{*}). \qquad (3.24)$$

Since  $\partial_{xx} \hat{v} > 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ , we have

$$\mathcal{J}_{\widehat{v},t}(-y-\mathcal{Y}_t)=x_t^*(y)=\widehat{x}^*=\mathcal{J}_{\widehat{v},t}(-Y_t^*-\mathcal{Y}_t).$$

Since  $\mathcal{J}_{\hat{v},t}(-y - \mathcal{Y}_t)$  is strictly decreasing with respect to y, we conclude that  $y = Y_t^*$ and  $Y^* = Y^{t,y;\tau^*,c^*,l^*,\pi^*}$ .

Finally, (3.23) and (3.24) imply that the value V takes the form of (3.19), and  $(\tau^*, c^*, l^*, \pi^*)$  is the optimal strategy. Moreover, by the continuity of  $\hat{v}_t(x), x_t^*(y)$  and (3.19), we derive V is continuous with respect to y, and

$$\begin{aligned} \partial_x \widehat{v}_t(x_t^*(y)) &= -y - \mathcal{Y}_t, \\ \partial_y V_t(y) &= \partial_x \widehat{v}_t(x_t^*(y)) \partial_y x_t^*(y) + \partial_y x_t^*(y)(y + \mathcal{Y}_t) + x_t^*(y) \\ &= x_t^*(y), \end{aligned}$$

which means that  $\partial_y V_t(y)$  is continuous with respect to y, too. And from (3.19), we compute the following second order partial derivative,

$$\partial_{yy}V_t(y) = \partial_y x_t^*(y) = \frac{-1}{\partial_{xx}\widehat{v}_t(x_t^*(y))} < 0$$

a.e. in  $\{(\omega, t, y) : y > -\mathcal{Y}_t(\omega), t \in [0, T], \omega \in \Omega\}$ . This completes the proof.  $\Box$ 

**Lemma 2** Suppose that the assumptions in Theorem 1 are satisfied. Then, the strategy  $\pi^*$  in Theorem 1 is well-defined and can be constructed from the solution to BSDE (3.20), and  $(\tau^*, c^*, l^*, \pi^*) \in \mathcal{A}^1(t, Y_t^*)$ .

**Proof** We prove the existence of  $\pi^*$ . In fact, SDE (3.9) implies that  $X^* \in S^p$  for any  $p \ge 1$  (refer to [23]). Denote  $\widehat{X}^* \triangleq 1/X^*$ , then it is not difficult to deduce that  $\widehat{X}^*$  is

governed by

$$\widehat{X}_s^* = \frac{1}{x^*} + \int_t^s \left( r_u - \beta + |\theta_u|^2 \right) \widehat{X}_u^* du + \int_t^s \widehat{X}_u^* \theta_u^\top (dW_u^\top, dB_u^\top)^\top, \quad \forall \ s \in [t, T].$$

Thus, we can claim that  $(X^*)^{-1} = \widehat{X}^* \in S^p$  for any  $p \ge 1$ . Hence, (2.2) implies that  $Y_T^* = \mathcal{J}_{U_2}(X_T^*) \in L^p(\mathcal{F}_T)$  for any  $p \ge 1$ . Repeating the same argument as in the above, we derive that  $c^*, l^* \in S^p$ . Consider the following BSDE,

$$Y_{s}^{*} = Y_{T}^{*} - \int_{s}^{T} \left[ Z_{u}^{\top} \theta_{u} + r_{u} Y_{u}^{*} - c_{u}^{*} + w_{u} (L_{u} - l_{u}^{*}) I_{\{u \leq \tau^{*}\}} \right] du$$
  
$$- \int_{s}^{T} Z_{u}^{\top} (dW_{u}^{\top}, dB_{u}^{\top})^{\top}.$$
(3.25)

It is clear that BSDE (3.25) has a unique solution  $(Y^*, Z) \in S^p \times \mathcal{L}^p$  for any  $p \ge 1$ . Since  $\Sigma$  is strongly non-degenerate, we can get  $\pi^* = (\Sigma^{-1})^T Z \in \mathcal{L}^p$  for any  $p \ge 1$ .

Next, we prove that  $(\tau^*, c^*, l^*, \pi^*) \in \mathcal{A}^1(t, Y_t^*)$ . In fact, from  $c^* \in \mathcal{S}^p$ ,  $\pi^* \in \mathcal{L}^p$  for any  $p \ge 1$ , we know that

$$\int_t^T (c_s + |\pi_s|^2) ds < \infty.$$

Moreover, since the ranges of the functions  $\mathcal{J}_{U_1}(\cdot; l)$ ,  $\mathcal{J}_{U_2}$  are  $(0, +\infty)$ , we deduce that  $Y_T^* > 0$ ,  $c^* > 0$ . And the definition of  $l^*$  implies that  $0 < l_s \leq \overline{L}_s \leq C$  for any  $t \leq s \leq \tau^*$ , and  $l_s = L_s \leq C$  for any  $\tau^* < s \leq T$ , where the constant *C* is the constant in Assumption 2.

Recalling the definitions of  $c^*$ ,  $l^*$  and  $Y_T^*$  and (3.3), (3.4), (3.5), we deduce that

$$\begin{aligned} U_1(c_s^*, l_s^*) &= \left[ X_s^*(c_s^* + w_s l_s^*) + \widehat{U}_{1,s}(X_s^*) \right] I_{\{t \le s \le \tau^*\}} \\ &+ \left[ X_s^* c_s^* + \widetilde{U}_{1,s}(X_s^*) \right] I_{\{\tau^* < s \le T\}} \\ &\geq - \left[ X_s^*(c_s^* + w_s l_s^*) + U_1(1, \overline{L}_s) - X_s^*(1 + \overline{L}_s w_s) \right]^- \\ &- \left[ X_s^* c_s^* + U_1(1, L_s) - X_s^* \right]^-; \\ U_2(Y_T^*) &= X_T^* Y_T^* + \widetilde{U}_2(X_T^*) \ge X_T^* Y_T^* + U_2(1) - X_T^*. \end{aligned}$$

Combining the above inequalities with the fact that  $X^*$ ,  $Y^*$ ,  $c^* \in S^p$  for any  $p \ge 1$  and w and  $l^*$  are bounded, we have

$$\mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} U_{1}^{-}(c_{s}^{*}, l_{s}^{*}) \, ds + e^{-\beta(T-t)} U_{2}^{-}(Y_{T}^{*})\right] < +\infty.$$

🖄 Springer

Moreover, it is not difficult to check that  $(Y^* + \mathcal{Y}(\cdot)I_{\{s < \tau^*\}}, \Sigma^{\top}\pi)$  with  $\pi = \pi^* - \pi^{\mathcal{Y}}I_{\{s < \tau^*\}}$  satisfies the following BSDE,

$$\begin{aligned} Y_{u}^{*} + \mathcal{Y}_{u}I_{\{s < \tau^{*}\}} &= Y_{T}^{*} - \int_{u}^{T} \Big[ \pi_{\xi}^{\top} \Sigma_{\xi} \theta_{\xi} + r_{\xi} (Y_{\xi}^{*} + \mathcal{Y}_{\xi} I_{\{\xi < \tau^{*}\}}) \\ &+ L_{\xi} w_{\xi} \left( I_{\{\xi \le \tau^{*}\}} - I_{\{s < \tau^{*}\}} \right) - c_{\xi}^{*} - w_{\xi} I_{\xi}^{*} I_{\{\xi \le \tau^{*}\}} \Big] d\xi \\ &- \int_{u}^{T} \pi_{\xi}^{\top} \Sigma_{\xi} \left( dW_{\xi}^{\top}, dB_{\xi}^{\top} \right)^{\top} \end{aligned}$$

for any  $t \le s \le u \le T$ . Since

$$Y_T^* > 0, \quad -L_{\xi} w_{\xi} \left( I_{\{\xi \le \tau^*\}} - I_{\{s < \tau^*\}} \right) + c_{\xi}^* + w_{\xi} l_{\xi}^* I_{\{\xi \le \tau^*\}} > 0,$$

applying the comparison theory for BSDEs, we deduce that  $Y_u^* + \mathcal{Y}_u I_{\{s < \tau^*\}} > 0$  for any  $t \le s \le u \le T$ . Particularly, setting u = s, we have that  $Y_s^* + \mathcal{Y}_s I_{\{s < \tau^*\}} > 0$  for any  $s \in [t, T]$ . Hence, we have proved that  $(\tau^*, c^*, l^*, \pi^*) \in \mathcal{A}^1(t, Y_t^*)$ .  $\Box$ 

#### 4 Verification of the Assumptions in Theorem 1

In this section, we establish the existence and uniqueness of the strong solutions to BSPDE(3.10) and BSPDVI(3.16), and show validity of the assumptions in Theorem 1.

#### 4.1 Transformation for Removing the Degenerateness

In order to remove the degenerateness of the operator  $\mathcal{L}$ , we introduce the following transformations

$$\begin{aligned} \widetilde{x} &= \log x, \quad \widehat{P}_t(\widetilde{x}) = \widehat{V}_t(x), \quad \widehat{Q}_t(\widetilde{x}) = \widehat{Z}_t(x), \quad \underline{\widehat{P}}_t(\widetilde{x}) = \underline{\widehat{V}}_t(x), \quad \underline{\widetilde{P}}_t(\widetilde{x}) = \underline{\widetilde{V}}_t(x), \\ \underline{\widetilde{Q}}_t(\widetilde{x}) &= \underline{\widetilde{Z}}_t(x), \quad \widetilde{R}_{1,t}(\widetilde{x}) = \overline{\widetilde{U}}_{1,t}(x), \quad \widetilde{R}_2(\widetilde{x}) = \overline{\widetilde{U}}_2(x), \quad \widehat{R}_{1,t}(\widetilde{x}) = \widehat{U}_{1,t}(x). \end{aligned}$$
(4.1)

It is clear that BSPDE (3.10) is equivalent to the following BSPDE,

$$\begin{cases} d\underline{\widetilde{P}}_{t} = -\left(\widetilde{\mathcal{L}}\underline{\widetilde{P}}_{t} + \sum_{i=1}^{N_{2}}\widetilde{\mathcal{M}}_{i}\underline{\widetilde{\mathcal{Q}}}_{i,t} + \widetilde{R}_{1,t}\right) dt \\ + \sum_{i=1}^{N_{2}}\underline{\widetilde{\mathcal{Q}}}_{i,t} dB_{i,t} \text{ in } \Omega \times [0,T] \times \mathbb{R}; \\ \underline{\widetilde{P}}_{T}(\widetilde{x}) = \widetilde{R}_{2}(\widetilde{x}) \text{ for any } \widetilde{x} \in \mathbb{R} \text{ a.s in } \Omega, \end{cases}$$

$$(4.2)$$

where

$$\widetilde{\mathcal{L}} \triangleq \frac{1}{2} |\theta|^2 \partial_{\widetilde{x}\widetilde{x}} + \left(\beta - r - \frac{1}{2} |\theta|^2\right) \partial_{\widetilde{x}} - \beta, \ \widetilde{\mathcal{M}}_i \triangleq -\theta_i^2 \partial_{\widetilde{x}}, \ i = 1, \cdots, N_2.$$
(4.3)

And BSPDVI (3.16) is equivalent to the following BSPDVI,

$$\begin{aligned} d\widehat{P}_{t} &= -\left(\widetilde{\mathcal{L}}\widehat{P}_{t} + \sum_{i=1}^{N_{2}}\widetilde{\mathcal{M}}_{i}\widehat{Q}_{i,t} + \widehat{R}_{1,t}\right) dt \\ &+ \sum_{i=1}^{N_{2}}\widehat{Q}_{i,t} dB_{i,t} & \text{if } \widehat{P} > \underline{\widehat{P}}; \\ d\widehat{P}_{t} &\leq -\left(\widetilde{\mathcal{L}}\widehat{P}_{t} + \sum_{i=1}^{N_{2}}\widetilde{\mathcal{M}}_{i}\widehat{Q}_{i,t} + \widehat{R}_{1,t}\right) dt \\ &+ \sum_{i=1}^{N_{2}}\widehat{Q}_{i,t} dB_{i,t} & \text{if } \widehat{P} = \underline{\widehat{P}}; \\ \widehat{P}_{T}(\widetilde{x}) &= \underline{\widehat{P}}_{T}(\widetilde{x}) \text{ for any } \widetilde{x} \in \mathbb{R} \text{ a.s in } \Omega, \end{aligned}$$

$$(4.4)$$

where

$$\underline{\widehat{P}}_{t}(\widetilde{x}) = \underline{\widehat{V}}_{t}(x) = \underline{\widetilde{V}}_{t}(x) - x\mathcal{Y}_{t} = \underline{\widetilde{P}}_{t}(\widetilde{x}) - e^{\widetilde{x}}\mathcal{Y}_{t}.$$

#### 4.2 Existence and Uniqueness of the Strong Solution to BSPDVI (3.16)

From (3.7), we know that there exist positive constants C and K such that

$$\begin{aligned} |\widetilde{R}_{1,t}(\widetilde{x})| &+ |\widetilde{R}_2(\widetilde{x})| + |\widetilde{R}'_2(\widetilde{x})| + |\widehat{R}_{1,t}(\widetilde{x})| \\ &\leq C \, e^{(K+1)|\widetilde{x}|}, \,\,\forall \, \widetilde{x} \in \mathbb{R} \text{ a.e. in } \Omega \times [0,T]. \end{aligned}$$
(4.5)

Hence, we deduce that  $\widetilde{R}_1$ ,  $\widehat{R}_1 \in \mathbb{H}^{0,2}_{K+2}$ , and  $\widetilde{R}_2 \in \mathbb{L}^{1,2}_{K+2}$ . Applying Lemmas 6 and 4, we can show the existence and uniqueness of the strong solutions to BSPDE (4.2) and BSPDVI (4.4).

**Theorem 2** BSPDE (4.2) has a unique strong solution  $(\underline{\widetilde{P}}, \underline{\widetilde{Q}}) \in \mathbb{H}^{2,2}_{K+2} \times \mathbb{H}^{1,2}_{K+2}$ , and BSPDVI (4.4) has a unique strong solution  $(\widehat{P}, \widehat{Q}, \widehat{k}^+) \in \mathbb{H}^{2,2}_{K+2} \times \mathbb{H}^{1,2}_{K+2} \times \mathbb{H}^{0,2}_{K+2}$ . Moreover,  $\underline{\widetilde{P}}, \ \widehat{P} \in \mathbb{S}^{1,2}_{K+2}$ .

**Proof** First, since  $\widetilde{R}_1 \in \mathbb{H}^{0,2}_{K+2}$ , and  $\widetilde{R}_2 \in \mathbb{L}^{1,2}_{K+2}$ , BSPDE (4.2) has a unique strong solution  $(\underline{\widetilde{P}}, \underline{\widetilde{Q}}) \in \mathbb{H}^{2,2}_{K+2} \times \mathbb{H}^{1,2}_{K+2}$ , and  $\underline{\widetilde{P}} \in \mathbb{S}^{1,2}_{K+2}$ . Next, we consider the properties of  $\underline{\widehat{P}}$ , which is the lower obstacle of BSPDVI (4.4). Since  $\underline{\widehat{P}}_t(\widetilde{x}) = \underline{\widetilde{P}}_t(\widetilde{x}) - e^{\widetilde{x}} \mathcal{Y}_t$ ,  $\mathcal{Y} \in S^2$  and  $\mathcal{Y}$  is  $\mathcal{P}^B$  – measurable, we know that  $\underline{\widehat{P}} \in \mathbb{H}^{2,2}_{K+2} \cap \mathbb{S}^{1,2}_{K+2}$ , and  $\underline{\widehat{P}}_T \in \mathbb{L}^{1,2}_{K+2}$ . Moreover, by BSPDE (4.2) and (2.5)<sup>8</sup>, we see that  $\underline{\widehat{P}}$  is a continuous semimartingale satisfying  $d\underline{\widehat{P}}_t = \underline{g}_t dt + \sum_{i=1}^{N_2} \underline{w}_{i,t} dB_{i,t}$ , where

$$\underline{g} = -\left[\widetilde{\mathcal{L}}\underline{\widetilde{P}} + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \underline{\widetilde{\mathcal{Q}}}_i + \widetilde{R}_1 + e^{\widetilde{x}} \left(\sum_{i=1}^{N_2} Z_{i+N_1}^{\mathcal{Y}} \theta_i^2 + r\mathcal{Y} - L w\right)\right],$$

<sup>8</sup> Note that  $Z_i^{\mathcal{Y}} = 0, i = 1, \cdots, N_1$ .

Applied Mathematics & Optimization

 $\underline{w}_i = \underline{\widetilde{Q}}_i - e^{\widetilde{x}} Z_{i+N_1}^{\mathcal{Y}}.$ 

Since  $Z^{\mathcal{Y}} \in \mathcal{L}^2$ , we deduce that  $\underline{g} \in \mathbb{H}^{0,2}_{K+2}$  and  $\underline{w} \in \mathbb{H}^{1,2}_{K+2}$ . Combining  $\widehat{R}_1 \in \mathbb{H}^{0,2}_{K+2}$ , and applying Lemma 4, we derive that BSPDVI (4.4) has a unique strong solution  $(\widehat{P}, \widehat{Q}, \widehat{K}^+) \in \mathbb{H}^{2,2}_{K+2} \times \mathbb{H}^{1,2}_{K+2} \times \mathbb{H}^{0,2}_{K+2}$ , and  $\widehat{P} \in \mathbb{S}^{1,2}_{K+2}$ .

From transformation (4.1), it is easy to deduce the following results.

**Theorem 3** *BSPDE* (3.10) *has a unique strong solution*  $(\underline{\widetilde{V}}, \underline{\widetilde{Z}})$ *, and BSPDVI* (3.16) *has a unique strong solution*  $(\widehat{V}, \widehat{Z}, \widehat{k}^+)$ *. Moreover,*  $\underline{\widetilde{V}}, \ \widehat{V} \in \mathbb{S}_1^{1,2}(D)$  *for any compact set D of*  $\mathbb{R}^+$ *.* 

## 4.3 Properties of the Strong Solution to BSPDVI (4.4)

In this subsection, we consider the properties of the strong solution to BSPDVI (3.16) via BSPDE (4.2) and BSPDVI (4.4), which constitute the assumptions in Theorem 1. First, we consider the growth properties of  $\underline{\widetilde{V}}$  and  $\widehat{V}$ .

**Theorem 4** There exists a constant C such that

$$|\underline{\widetilde{P}}_{t}(\widetilde{x})|, |\widehat{P}_{t}(\widetilde{x})| \leq C e^{(K+1)|\widetilde{x}|}, \quad \forall \, \widetilde{x} \in \mathbb{R} \text{ a.e. in } \Omega \times [0, T].$$

$$(4.6)$$

Hence,  $\underline{\widetilde{V}}$  and  $\widehat{V}$  have the following growth properties

$$|\underline{\widetilde{V}}_t(\widetilde{x})| + |\widehat{V}_t(\widetilde{x})| \le C \left( x^{K+1} + x^{-(K+1)} \right), \quad \forall x \in \mathbb{R}^+ \ a.e. \ in \ \Omega \times [0, T].$$

**Proof** It is sufficient to prove (4.6), and the estimates about  $\underline{\tilde{V}}$  and  $\hat{V}$  follow from transformation (4.1). In order to prove (4.6), we first introduce an auxiliary function, which will be used repeatedly below,

$$\phi(t, \widetilde{x}; k_1, k_2) = e^{2k_1(T-t)} \left( e^{k_2 |\widetilde{x}|} - k_2 |\widetilde{x}| - 1 \right) \ge 0, \tag{4.7}$$

where  $k_1, k_2$  are non-negative numbers. Since  $e^x - x > e^x/2$ , we can show that  $\phi \in C^2$ , and

$$\begin{split} \phi(t,0;k_{1},k_{2}) &= \partial_{\widetilde{x}}\phi(t,0;k_{1},k_{2}) = 0; \\ &-\partial_{t}\phi = 2k_{1}\phi \geq k_{1}e^{2k_{1}(T-t)}\left(e^{k_{2}|\widetilde{x}|} - 2\right); \\ &|\partial_{\widetilde{x}}\phi| = k_{2}e^{2k_{1}(T-t)}\left(e^{k_{2}|\widetilde{x}|} - 1\right) \leq k_{2}e^{2k_{1}(T-t)}e^{k_{2}|\widetilde{x}|}; \\ &\partial_{\widetilde{x}\widetilde{x}}\phi = k_{2}^{2}e^{2k_{1}(T-t)}e^{k_{2}|\widetilde{x}|}; \\ &-\partial_{t}\phi - \widetilde{\mathcal{L}}\phi \geq e^{2k_{1}(T-t)}\left\{e^{k_{2}|\widetilde{x}|}\left[k_{1} - \frac{k_{2}^{2}|\theta_{t}|^{2}}{2} - \left(\beta + |r_{t}| + \frac{|\theta_{t}|^{2}}{2}\right)k_{2}\right] - 2k_{1}\right\}. \end{split}$$

$$(4.8)$$

🖄 Springer

Construct a super-solution to BSPDE (4.2) as the following,

$$\overline{P}_t(\widetilde{x}) = 2C\phi(t,\widetilde{x};M,K+1) + 2Ce^{2M(T-t)}, \quad \overline{Q} = 0,$$
(4.9)

where C, K are the constants in (4.5), and M is a positive constant which will be determined later. From (4.8), (4.5) and Assumption 2, we deduce that

$$\begin{split} \overline{R}_t &\triangleq -\partial_t \overline{P}_t - \widetilde{\mathcal{L}} \overline{P}_t \\ &\geq 2C e^{2M(T-t) + (K+1)|\widetilde{x}|} \left[ M - \frac{(K+1)^2 |\theta_t|^2}{2} - \left(\beta + |r_t| + \frac{|\theta_t|^2}{2}\right) (K+1) \right] \\ &\geq C e^{(K+1)|\widetilde{x}|} \\ &\geq \widetilde{R}_1, \end{split}$$

provided *M* is large enough. It is clear that  $(\overline{P}, \overline{Q})$  satisfies the following BSPDE

$$\begin{cases} d\overline{P}_t = -\left(\widetilde{\mathcal{L}P}_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \overline{\mathcal{Q}}_{i,t} + \overline{R}_t\right) dt + \sum_{i=1}^{N_2} \overline{\mathcal{Q}}_{i,t} dB_{i,t} \text{ in } \Omega \times [0, T] \times \mathbb{R};\\ \overline{P}_t(\widetilde{x}) > Ce^{(K+1)|\widetilde{x}|} \ge R_2(\widetilde{x}) \text{ for any } \widetilde{x} \in \mathbb{R} \text{ a.s. in } \Omega. \end{cases}$$

Recalling Lemma 7, we deduce that

$$\underline{\widetilde{P}} \leq \overline{P} \leq 2Ce^{2M(T-t)} \Big( e^{(K+1)|\widetilde{x}|} - (K+1)|\widetilde{x}| \Big).$$

Repeating the same argument as above, we can prove that

$$\underline{\widetilde{P}} \ge -\overline{P} \ge -2Ce^{2M(T-t)} \Big( e^{(K+1)|\widetilde{x}|} - (K+1)|\widetilde{x}| \Big).$$

Then the conclusion about  $\underline{\widetilde{P}}$  is obvious if we change the constant C.

Recalling (2.8), we have the following estimate about the lower obstacle  $\underline{\widehat{P}}$ ,

$$\left|\underline{\widehat{P}}_{t}(\widetilde{x})\right| = \left|\underline{\widetilde{P}}_{t}(\widetilde{x}) - e^{\widetilde{x}}\mathcal{Y}_{t}\right| \le Ce^{(K+1)|\widetilde{x}|},\tag{4.10}$$

where we have changed the constant *C*. Construct an auxiliary function as (4.9), where *C* is large enough such that  $\overline{P} > \underline{\widehat{P}}$ . Then  $(\overline{P}, \overline{Q}, 0), \overline{Q} = 0$  satisfies the following BSPDVI,

$$\begin{cases} d\overline{P}_t = -\left(\widetilde{\mathcal{L}P}_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}_i \overline{Q}_{i,t}} + \overline{R}_t\right) dt + \sum_{i=1}^{N_2} \overline{Q}_{i,t} dB_{i,t} & \text{if } \overline{P} > \underline{\widehat{P}}; \\ d\overline{P}_t \leq -\left(\widetilde{\mathcal{L}P}_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \overline{Q}_{i,t} + \overline{R}_t\right) dt + \sum_{i=1}^{N_2} \overline{Q}_{i,t} dB_{i,t} & \text{if } \overline{P} = \underline{\widehat{P}}; \\ \overline{P}_T(\widetilde{x}) > \underline{\widehat{P}}_T(\widetilde{x}). \end{cases}$$

🖄 Springer

Since  $\overline{R} \ge \widehat{R}_1$ , we see that  $\widehat{P} \le \overline{P}$  by Lemma 5. So, we can deduce that

$$-Ce^{(K+1)|\widetilde{x}|} \leq \underline{\widehat{P}} \leq \widehat{P} \leq \overline{P} \leq Ce^{(K+1)|\widetilde{x}|}.$$

Until now, we have proved all conclusions in this theorem.

Since the theory of BSPDEs is not as mature as the theory of PDEs, we need to utilize the corresponding optimal stopping problem to analyze the other properties of  $\hat{V}$ . We first give the relationship between BSPDVI (3.16) and the corresponding optimal stopping problem by the following theorem.

**Theorem 5** Let  $(\underline{\tilde{V}}, \underline{\tilde{Z}})$  and  $(\widehat{V}, \widehat{Z}, \widehat{k}^+)$  be the strong solutions to BSPDE (3.10) and BSPDVI (3.16), respectively. Then

$$\underline{\widetilde{V}}_{t}(x) = \mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} \widetilde{U}_{1}(X_{s}^{t,x}) ds + e^{-\beta(T-t)} \widetilde{U}_{2}(X_{T}^{t,x}) \left| \mathcal{F}_{t} \right.\right], \quad (4.11)$$

and  $\widehat{V}_t(x)$  is the value of the following optimal stopping problem,

$$\widehat{V}_{t}(x) = \underset{\tau \in \mathcal{U}_{t,T}}{\operatorname{ess.sup}} \mathbb{E}\left[\int_{t}^{\tau \wedge T} e^{-\beta(s-t)} \widehat{U}_{1,s}(X_{s}^{t,x}) ds + e^{-\beta(\tau-t)} \underline{\widehat{V}}_{\tau \wedge T}(X_{\tau \wedge T}^{t,x}) \left| \mathcal{F}_{t} \right. \right],$$

$$(4.12)$$

and its optimal stopping time  $\tau^*$  can be described as

$$\tau^* \triangleq \inf \left\{ s \in [t, T] : \widehat{V}_s(X_s^{t, x}) = \underline{\widehat{V}}_s(X_s^{t, x}) \right\} \wedge T.$$

**Proof** From (4.6) and  $\widetilde{X} = \log X$ , we deduce that

$$\left|\frac{\widetilde{P}}{P}\left(\widetilde{X}^{t,\widetilde{x}}\right)\right|, \left|\widehat{P}\left(\widetilde{X}^{t,\widetilde{x}}\right)\right| \leq C e^{(K+1)|\widetilde{X}^{t,\widetilde{x}}|} \leq C \left(\left|X^{t,x}\right|^{K+1} + \left|X^{t,x}\right|^{-K-1}\right).$$

Recalling (4.5), and  $X^{-1}$ ,  $X \in S^p$  for any p > 1, we know that

$$\widehat{R}_1(\widetilde{X}^{t,\widetilde{x}}), \ \underline{\widehat{P}}(\widetilde{X}^{t,\widetilde{x}}), \ \widehat{P}(\widetilde{X}^{t,\widetilde{x}}) \in \mathcal{S}^2.$$

So, we have shown all assumptions in Lemma 8 are valid, and we can derive the conclusions in Lemma 8. Then the conclusions in this theorem follow from transformation (4.1).

Next, we utilize (4.11) and (4.12) to analyze the properties of  $\widehat{V}$ .

**Theorem 6** Let  $(\underline{\widetilde{V}}, \underline{\widetilde{Z}})$  and  $(\widehat{V}, \widehat{Z}, \widehat{k}^+)$  be the strong solutions to BSPDE (3.10) and BSPDVI (3.16), respectively. Then  $\underline{\widetilde{V}}$  and  $\widehat{V}$  are strictly concave, i.e.,  $\partial_{xx}\underline{\widetilde{V}}, \ \partial_{xx}\widehat{V} > 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ .

**Proof** Let  $x_1, x_2 \in \mathbb{R}^+$ ,  $0 < \lambda < 1$ ,  $x_{\lambda} = \lambda x_1 + (1 - \lambda) x_2$ . From (3.9), it is clear that

$$X^{t,x_{\lambda}} = \lambda X^{t,x_1} + (1-\lambda)X^{t,x_2}.$$

Since  $\widetilde{U}_1$  and  $\widetilde{U}_2$  are strictly convex, we have

$$\begin{aligned} \widetilde{U}_1(X^{t,x_{\lambda}}) &< \lambda \widetilde{U}_1(X^{t,x_1}) + (1-\lambda) \widetilde{U}_1(X^{t,x_2}), \\ \widetilde{U}_2(X_T^{t,x_{\lambda}}) &< \lambda \widetilde{U}_2(X_T^{t,x_1}) + (1-\lambda) \widetilde{U}_2(X_T^{t,x_2}). \end{aligned}$$

So, (4.11) implies that  $\underline{\widetilde{V}}_t(x_\lambda) < \lambda \underline{\widetilde{V}}_t(x_1) + (1-\lambda)\underline{\widetilde{V}}_t(x_2)$ . Combining  $\underline{\widetilde{V}} \in \mathbb{H}^{2,2}(D)$  for any compact subset D of  $\mathbb{R}$ , we see that  $\partial_{xx}\underline{\widetilde{V}}_t(x) > 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ . Since  $\underline{\widetilde{V}}_t(x) = \underline{\widetilde{V}}_t(x) - x\mathcal{Y}_t$ , we derive that

$$\begin{aligned} \widehat{\underline{V}}_{t}(x_{\lambda}) &= \widetilde{\underline{V}}_{t}(x_{\lambda}) - x_{\lambda} \mathcal{Y}_{t} \\ &< \lambda \widetilde{\underline{V}}_{t}(x_{1}) + (1 - \lambda) \widetilde{\underline{V}}_{t}(x_{2}) - x_{\lambda} \mathcal{Y}_{t} \\ &= \lambda \widehat{\underline{V}}_{t}(x_{1}) + (1 - \lambda) \widetilde{\underline{V}}_{t}(x_{2}). \end{aligned}$$

Denote

$$\tau^{\lambda} = \inf\left\{s \in [t, T] : \widehat{V}_{s}(X_{s}^{t, x_{\lambda}}) = \underline{\widehat{V}}_{s}(X_{s}^{t, x_{\lambda}})\right\} \wedge T,$$
(4.13)

then we have

$$\begin{split} \widehat{V}_{t}(x_{\lambda}) &= \mathbb{E}\left[\int_{t}^{\tau^{\lambda}} e^{-\beta(s-t)} \widehat{U}_{1,s}(X_{s}^{t,x_{\lambda}}) ds + e^{-\beta(\tau^{\lambda}-t)} \underline{\widehat{V}}_{\tau^{\lambda}}(X_{\tau^{\lambda}}^{t,x_{\lambda}}) \left| \mathcal{F}_{t} \right. \right] \\ &< \mathbb{E}\left[\int_{t}^{\tau^{\lambda}} e^{-\beta(s-t)} \Big[ \lambda \widehat{U}_{1,s}(X_{s}^{t,x_{1}}) + (1-\lambda) \widehat{U}_{1,s}(X_{s}^{t,x_{2}}) \right] ds \\ &+ e^{-\beta(\tau^{\lambda}-t)} \Big[ \lambda \underline{\widehat{V}}_{\tau^{\lambda}}(X_{\tau^{\lambda}}^{t,x_{1}}) + (1-\lambda) \underline{\widehat{V}}_{\tau^{\lambda}}(X_{\tau^{\lambda}}^{t,x_{2}}) \Big] \left| \mathcal{F}_{t} \right. \Big] \\ &\leq \lambda \widehat{V}_{t}(x_{1}) + (1-\lambda) \widehat{V}_{t}(x_{2}). \end{split}$$

Combining  $\widehat{V} \in \mathbb{H}^{2,2}(D)$  for any compact subset D of  $\mathbb{R}$ , we know that  $\partial_{xx} \widehat{V}_t(x) > 0$ a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ . 

**Theorem 7**  $\partial_x \underline{\widetilde{V}}_t(x), \partial_x \widehat{V}_t(x) \to -\infty \text{ as } x \to 0^+, \text{ and }$ 

$$\partial_x \underline{\widetilde{V}}_t(x), \partial_x \widehat{V}_t(x) \to 0 \text{ as } x \to +\infty \text{ a.e. in } \Omega \times [0, T].$$

**Proof** Denote  $\widetilde{H}_s = e^{\beta(s-t)}H_s^t \ge H_s^t > 0$ , then  $X^{t,x} = x \widetilde{H}$ . Let  $x \in \mathbb{R}^+$  and  $1/2 < \lambda < 1$ . Since  $\widetilde{U}_i$  is strictly convex and decreasing with respect to x, we have

$$\partial_x \widetilde{U}_i \left( X^{t,\lambda x} \right) (1-\lambda) X \widetilde{H} \le \widetilde{U}_i (X^{t,x}) - \widetilde{U}_i (X^{t,\lambda x})$$

$$\leq \partial_x \widetilde{U}_i(X^{t,x})(1-\lambda)x\widetilde{H}$$
  
$$\leq 0, \qquad i = 1, 2. \tag{4.14}$$

Then, by (4.11), we deduce that

$$0 \geq \frac{\widetilde{\underline{V}}_{t}(x) - \widetilde{\underline{V}}_{t}(\lambda x)}{(1 - \lambda)x}$$
  
$$\geq \mathbb{E}\left[\int_{t}^{T} \partial_{x} \widetilde{U}_{1,s}\left(X_{s}^{t,\lambda x}\right) H_{s}^{t} ds + \widetilde{U}_{2}^{\prime}\left(X_{T}^{t,\lambda x}\right) H_{T}^{t} \middle| \mathcal{F}_{t} \right].$$
(4.15)

Next, we estimate the two terms inside the braces. From Lemma 1, we see that for any positive number  $\varepsilon$ , there exists a positive constant  $M_{\varepsilon} > 1$  such that

$$\partial_x \widetilde{U}_{1,s}(z) > -\varepsilon, \ \forall \ z \ge M_{\varepsilon}; \quad \partial_x \widetilde{U}_{1,s}(z) > \partial_x \widetilde{U}_{1,s}(1) - C(z^{-K} + 1), \ \forall \ z > 0,$$

a.e. in  $\Omega \times [t, T]$ . So, we have

$$\begin{split} & \mathbb{E}\left[\int_{t}^{T} \partial_{x} \widetilde{U}_{1,s}\left(X_{s}^{t,\lambda x}\right) H_{s}^{t} ds \left|\mathcal{F}_{t}\right.\right] \\ & \geq \mathbb{E}\left[\int_{t}^{T} \partial_{x} \widetilde{U}_{1,s}\left(\frac{1}{2} x \widetilde{H}_{s}\right) H_{s}^{t} ds \left|\mathcal{F}_{t}\right.\right] \\ & \geq \mathbb{E}\left[-\varepsilon \int_{t}^{T} H_{s}^{t} I_{\{x \widetilde{H}_{s} > 2M_{\varepsilon}\}} ds \right. \\ & \left.+\int_{t}^{T} \left[\partial_{x} \widetilde{U}_{1,s}(1) - C - C\left(\frac{1}{2} x \widetilde{H}_{s}\right)^{-K}\right] H_{s}^{t} I_{\{x \widetilde{H}_{s} \leq 2M_{\varepsilon}\}} ds \left|\mathcal{F}_{t}\right.\right] \\ & \geq -\varepsilon \mathbb{E}\left[\int_{t}^{T} H_{s}^{t} ds \left|\mathcal{F}_{t}\right.\right] + \left(\partial_{x} \widetilde{U}_{1,s}(1) - C\right) \\ & \times \mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} \frac{2M_{\varepsilon}}{x} I_{\{\widetilde{H}_{s} \leq \frac{2M_{\varepsilon}}{x}\}} ds \left|\mathcal{F}_{t}\right.\right] - \frac{C2^{K}}{x^{K}} \mathbb{E}\left[\int_{t}^{T} (H_{s}^{t})^{1-K} ds \left|\mathcal{F}_{t}\right.\right]. \end{split}$$

From (2.7) and Assumption 2, it is not difficult to deduce that

$$\mathbb{E}\left[\sup_{s\in[t,T]} (H_s^t)^p \left| \mathcal{F}_t \right.\right] < C_p, \quad \forall \ p > 1.$$
(4.16)

So, we can find a sufficiently large positive constant  $\widetilde{M}_{\varepsilon}$  independent of  $\lambda$  and t such that

$$\mathbb{E}\left[\int_{t}^{T}\partial_{x}\widetilde{U}_{1,s}\left(\frac{1}{2}x\widetilde{H}_{s}\right)H_{s}^{t}ds\left|\mathcal{F}_{t}\right]\geq-C\varepsilon,\quad\forall\ x\geq\widetilde{M}_{\varepsilon},$$

D Springer

where C is another positive constant independent of  $\varepsilon$ ,  $\widetilde{M}_{\varepsilon}$ , x,  $\lambda$ , t. Repeating the same argument as above, we can deduce

$$\mathbb{E}\left[\left.\widetilde{U}_{2}^{\prime}\left(\frac{1}{2}x\widetilde{H}_{T}\right)H_{T}^{t}\right|\mathcal{F}_{t}\right]\geq-C\varepsilon,\quad\forall\ x\geq\widetilde{M}_{\varepsilon}.$$
(4.17)

Hence, by (4.15), we have proved that

$$0 \geq \frac{\widetilde{V}_t(x) - \widetilde{V}_t(\lambda x)}{(1 - \lambda)x} \geq -2C\varepsilon, \quad \forall x \geq \widetilde{M}_{\varepsilon}.$$

Let  $\lambda \to 1^-$ , then we derive that  $\partial_x \underline{\widetilde{V}} \leq 0$ , and  $\partial_x \underline{\widetilde{V}}_t(x) \geq -2C\varepsilon$  for any  $x \geq \widetilde{M}_{\varepsilon}$ , where we have used the fact that  $\underline{\widetilde{V}} \in \mathbb{H}^{2,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ . By taking the limits  $x \to +\infty$  and  $\varepsilon \to 0^+$  sequentially, we deduce that ess.inf  $\{\partial_x \underline{\widetilde{V}}_t(x) : (\omega, t) \in \Omega \times [0, T]\} \to 0$  as  $x \to +\infty$ .

From Lemma 1, we see that, for any positive number M, there exists a positive constant  $\varepsilon_M$  such that

$$\widetilde{U}_2'(z) < -M, \ \forall \ 0 < z \le \varepsilon_M.$$

So, by (4.14) and some computation we show that

$$\begin{aligned} \frac{\widetilde{\underline{V}}_{t}(x) - \widetilde{\underline{V}}_{t}(\lambda x)}{(1 - \lambda)x} &\leq \mathbb{E}\left[\int_{t}^{T} \partial_{x} \widetilde{U}_{1,s}\left(x\widetilde{H}_{s}\right) H_{s}^{t} ds + \widetilde{U}_{2}^{\prime}\left(x\widetilde{H}_{T}\right) H_{T}^{t} \left|\mathcal{F}_{t}\right]\right] \\ &\leq -M\mathbb{E}\left[H_{T}^{t} I_{\{0 < x\widetilde{H}_{T} \leq \varepsilon_{M}\}} \left|\mathcal{F}_{t}\right]\right] \\ &\leq -M\mathbb{E}\left[H_{T}^{t} I_{\{0 < H_{T}^{t} \leq \frac{\varepsilon_{M}}{x} e^{-\beta T}\}} ds \left|\mathcal{F}_{t}\right]\right] \\ &\leq -\frac{M}{C}, \end{aligned}$$

provided  $0 < x < e^{-\beta T} \varepsilon_M$ , where *C* is a positive constant independent of  $M, \varepsilon_M, x, \lambda, t$ . By letting  $\lambda \to 1^-$ , we derive that  $\partial_x \underline{\widetilde{V}}_t(x) \leq -M/C$  for any  $0 < x < e^{-\beta T} \varepsilon_M$ . By taking limits  $x \to 0^+$  and  $M \to +\infty$  sequentially, we deduce that ess.sup{ $\partial_x \underline{\widetilde{V}}_t(x) : (\omega, t) \in \Omega \times [0, T]$ }  $\to -\infty$  as  $x \to 0^+$ .

Next, we prove that  $\widehat{V}$  has the same asymptotic properties. As in the above, we show by computation

$$0 \geq \frac{\widehat{V}_{t}(x) - \widehat{V}_{t}(\lambda x)}{(1 - \lambda)x}$$
  
$$\geq \mathbb{E}\left[\int_{t}^{\tau^{\lambda}} \partial_{x} \widehat{U}_{1,s}\left(X_{s}^{t,\lambda x}\right) H_{s}^{t} ds + \partial_{x} \underline{\widehat{V}}_{\tau^{\lambda}}\left(X_{\tau^{\lambda}}^{t,\lambda x}\right) H_{\tau^{\lambda}}^{t} \middle| \mathcal{F}_{t} \right]$$

🖄 Springer

$$\geq \mathbb{E}\left[\int_{t}^{T} \partial_{x}\widehat{U}_{1,s}\left(\frac{1}{2}x\widetilde{H}_{s}\right)H_{s}^{t}ds + \partial_{x}\underline{\widehat{V}}_{\tau^{\lambda}}\left(X_{\tau^{\lambda}}^{t,\lambda x}\right)H_{\tau^{\lambda}}^{t}\left|\mathcal{F}_{t}\right],\qquad(4.18)$$

where  $\tau^{\lambda}$  is the optimal stopping time with the initial state  $\lambda x$ , defined in (4.13). Next, we estimate the two terms inside the braces as in the above. It is clear that the first term is similar to the first term inside the braces in (4.15). So we can obtain its estimate by the same argument. Now, we focus on estimating the second term.

In fact, from Theorem 11, we conclude that  $\widehat{V}_s(X_s^{t,\lambda X}) > \widehat{V}_s(X_s^{t,\lambda X})$  for any  $X_s^{t,\lambda X} > \overline{X}$  and  $0 \le t < T$ , and  $\widehat{V}_s(X_s^{t,\lambda X}) = \widehat{V}_s(X_s^{t,\lambda X})$  for any  $X_s^{t,\lambda X} \le \underline{X}$ , where  $\overline{X}$  and  $\underline{X}$  are positive constants. So we deduce that  $\tau^{\lambda} = T$  in the event  $A = \{\omega : \inf\{X_s^{t,\lambda X}(\omega) : s \in [t, T]\} > \overline{X}\}$ , and  $X_{\tau^{\lambda}}^{t,\lambda X} \ge \underline{X}$  provided  $\lambda X \ge \underline{X}$ . Thus, if  $\lambda X \ge \underline{X}$ , we have

$$\mathbb{E}\left[\partial_{x} \underline{\widehat{V}}_{\tau^{\lambda}}\left(X_{\tau^{\lambda}}^{t,\lambda x}\right) H_{\tau^{\lambda}}^{t}\right] \\
= \mathbb{E}\left[\partial_{x} \underline{\widehat{V}}_{\tau^{\lambda}}\left(X_{\tau^{\lambda}}^{t,\lambda x}\right) H_{\tau^{\lambda}}^{t} I_{\Omega \setminus A} + \partial_{x} \underline{\widehat{V}}_{T}\left(X_{T}^{t,\lambda x}\right) H_{T}^{t} I_{A}\right] \\
\geq \mathbb{E}\left[\inf_{s \in [t,T]} \partial_{x} \underline{\widehat{V}}_{s}\left(\underline{X}\right) \sup_{s \in [t,T]} H_{s}^{t} I_{\Omega \setminus A} + \widetilde{U}_{2}^{\prime}\left(\frac{1}{2}x \widetilde{H}_{T}\right) H_{T}^{t}\right], \quad (4.19)$$

where we have used the facts that  $\partial_x \underline{\widehat{V}} = \partial_x \underline{\widetilde{V}} - \mathcal{Y}$  is increasing with respect to x and

$$X^{t,\lambda x} \ge X^{t,x/2} = \frac{1}{2}x\widetilde{H}, \quad \underline{\widehat{V}}_T(x) = \underline{\widetilde{V}}_T(x) - x\mathcal{Y}_T = \widetilde{U}_2(x) - x\mathcal{Y}_T, \quad \mathcal{Y}_T = 0.$$

It is clear that the second term on the right-hand side of (4.19) is the same as that in (4.15), and we can deduce the estimate (4.17).

Recalling (4.10) and  $\underline{\widetilde{V}} \in \mathbb{S}^{1,2}(D)$  for any compact subset of  $\mathbb{R}^+$ , we have

$$\mathbb{E}\left[\sup_{s\in[t,T]}\left|\partial_{x}\underline{\widehat{\mathcal{V}}}_{s}\left(\underline{X}\right)\right|^{2}\right] \leq \frac{4}{\underline{X}}\mathbb{E}\left[\sup_{s\in[t,T]}\int_{\underline{X}/2}^{\underline{X}}\left|\partial_{x}\underline{\widetilde{\mathcal{V}}}_{s}(x)\right|^{2}\right] + 2\mathbb{E}\left[\sup_{s\in[t,T]}\mathcal{Y}_{s}^{2}\right]$$
$$\leq C,$$

where we have used the fact that  $\partial_x \underline{\widetilde{V}}_s(\underline{X}) \leq \partial_x \underline{\widetilde{V}}_s(x) \leq 0$  for any  $x \in [\underline{X}/2, \underline{X}]$ . Moreover,

$$\Omega \setminus A = \left\{ \omega : \sup_{s \in [t,T]} \frac{1}{\lambda x \widetilde{H}_s(\omega)} \ge \frac{1}{\overline{X}} \right\} \subset \left\{ \omega : \sup_{s \in [t,T]} \frac{1}{H_s^t(\omega)} \ge \frac{x}{2\overline{X}} \right\}.$$

Since  $1/H^t \in S^p$  for any p > 1, we derive that  $\mathbb{P}(\Omega \setminus A) \to 0$  as  $x \to +\infty$ . Combining (4.16), we have the following estimate of the first term on the right-hand

🖉 Springer

side of (4.19),

$$0 \geq \mathbb{E}\left[\inf_{s\in[t,T]}\partial_{x}\widehat{\underline{V}}_{s}\left(\underline{X}\right)\sup_{s\in[t,T]}H_{s}^{t}I_{\Omega\setminus A}\right]$$
  
$$\geq -\mathbb{E}\left[\sup_{s\in[t,T]}\left|\partial_{x}\widehat{\underline{V}}_{s}\left(\underline{X}\right)\right|^{2}I_{\Omega\setminus A}\right]^{1/2}\mathbb{E}\left[\sup_{s\in[t,T]}\left(H_{s}^{t}\right)^{2}I_{\Omega\setminus A}\right]^{1/2} \to 0,$$
  
as  $x \to +\infty.$ 

So, from (4.18), we have proved that for any  $\varepsilon$ , there exists a positive constants  $M_{\varepsilon}$  such that

$$\frac{\widehat{V}_{t}(x)-\widehat{V}_{t}(\lambda x)}{(1-\lambda)x}\leq 0, \quad \mathbb{E}\left(\frac{\widehat{V}_{t}(x)-\widehat{V}_{t}(\lambda x)}{(1-\lambda)x}\right)\geq -2C\varepsilon, \quad \forall x\geq M_{\varepsilon}.$$

By letting  $\lambda \to 1^-$ , we derive that  $\partial_x \widehat{V} \leq 0$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}^+$ , and  $\mathbb{E}[\partial_x \widehat{V}_t(x)] \geq -2C\varepsilon$  for any  $x \geq \widetilde{M}_{\varepsilon}$ , where we have used the fact that  $\widehat{V} \in \mathbb{H}^{2,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ . By taking limits  $x \to +\infty$  and  $\varepsilon \to 0^+$  sequentially, we deduce that  $\mathbb{E}(\partial_x \widehat{V}_t(x)) \to 0$  as  $x \to +\infty$ . Combining the fact  $\partial_x \widehat{V}_t(x)$  is convex and less than zero, we conclue that  $\partial_x \widehat{V}_t(x) \to 0$  as  $x \to +\infty$  a.e. in  $\Omega \times [0, T]$ .

The proof of  $\partial_x \widehat{V}_t(x) \to -\infty$  as  $x \to 0^+$  is similar to the one in the above.  $\Box$ 

Until now, we have showed all the assumptions in Theorem 1 are satisfied. So, we have the following result.

**Theorem 8** The optimal retirement problem has a unique value V, which takes the form of (3.19). Moreover, the optimal investment, consumption and retirement strategy is described by the strategy described in Theorem 1.

#### **5** Properties of the Optimal Retirement Boundary

In this section, we utilize BSPDE (4.2) and BSPDVI (4.4) to study the properties of the optimal retirement boundary. For this purpose, we first investigate properties of the following functions,

$$\Delta P = \widehat{P} - \underline{\widehat{P}} = \widehat{P} - \underline{\widetilde{P}} + e^{\widetilde{x}} \mathcal{Y},$$
  

$$\Delta Q_i = \widehat{Q} - \underline{\widetilde{Q}}_i + e^{\widetilde{x}} Z_{N_1+i}^{\mathcal{Y}}, \qquad i = 1, 2, \cdots, N_2.$$
(5.1)

Recalling BSPDE (4.2), BSPDVI (4.4) and (2.5), we deduce  $\Delta P$  and  $\Delta Q$  satisfy the following BSPDVI,

$$\begin{cases} d\Delta P_t = -\left(\widetilde{\mathcal{L}}\Delta P_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \Delta Q_{i,t} + \Delta R_t\right) dt \\ + \sum_{i=1}^{N_2} \Delta Q_{i,t} dB_{i,t} & \text{if } \Delta P > 0; \\ d\Delta P_t \le -\left(\widetilde{\mathcal{L}}\Delta P_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \Delta Q_{i,t} + \Delta R_t\right) dt \\ + \sum_{i=1}^{N_2} \Delta Q_{i,t} dB_{i,t} & \text{if } \Delta P = 0; \\ \Delta P_T(\widetilde{x}) = 0, \end{cases}$$

$$(5.2)$$

where

$$\Delta R_t(\widetilde{x}) = \widehat{R}_{1,t}(\widetilde{x}) - \widetilde{R}_{1,t}(\widetilde{x}) + L_t w_t e^{\widetilde{x}} = \widehat{U}_{1,t}(x) - \widetilde{U}_{1,t}(x) + L_t w_t x.$$
(5.3)

From (4.5), we deduce that

$$|\Delta R_t(\widetilde{x})| \le C e^{(K+1)|x|}, \quad \forall \ \widetilde{x} \in \mathbb{R} \text{ a.e. in } \Omega \times [0, T].$$
(5.4)

In BSPDVI (5.2), the lower obstacle and the terminal value become 0, thus we have transformed the problem into a problem where the continuation value after retirement is 0.

**Theorem 9** There exist two constants  $\overline{X}_P$ ,  $\underline{X}_P$  such that  $\Delta P_t(\widetilde{x}) > 0$  if  $\widetilde{x} > \overline{X}_P$  and  $0 \le t < T$ , and  $\Delta P_t(\widetilde{x}) = 0$  if  $\widetilde{x} \le \underline{X}_P$ .

**Proof** We first prove that there exists  $\overline{X}_P$  such that  $\Delta P_t(\widetilde{x}) > 0$  if  $\widetilde{x} > \overline{X}_P$  and t < T. In order to show this property, we construct a function  $\underline{\Delta P}$  such that  $\underline{\Delta P} \leq \Delta P$  and  $\underline{\Delta P}_t(\widetilde{x}) > 0$  if  $\widetilde{x} > \overline{X}_P$  and  $0 \leq t < T$ .

From Lemma 10, there exist two positive constants  $\epsilon$  and  $\widetilde{X}_P$  such that  $\Delta R_t(\widetilde{x}) > \epsilon e^{\widetilde{x}}$  for any  $\widetilde{x} \geq \widetilde{X}_P$ . Construct  $\Delta P$  as the following,

$$\underline{\Delta P}_{t}(\widetilde{x}) = \frac{\epsilon}{1 + \|r\|_{\infty}} \left(1 - e^{t-T}\right) \phi(t, \widetilde{x} - \overline{X}_{P}; 0, 1) I_{\{\widetilde{x} > \overline{X}_{P}\}} \\ - C \phi(t, \widetilde{x} - \overline{X}_{P}; M, K + 1) I_{\{\widetilde{x} \le \overline{X}_{P}\}},$$

where the definition of  $||r||_{\infty}$  is the same as in (2.8), and  $\phi$  is defined in (4.7), *C*, *K* are the constants in (5.4), and M > 0,  $\overline{X}_P > 2\widetilde{X}_P$  will be determined later. Recalling the properties of  $\phi$ , we can check that  $\Delta P_T \leq 0$ , and  $\partial_{\widetilde{x}} \Delta P$  is locally Lipschitz-continuous with respect to *x* and smooth with respect to *t*, and  $\Delta P \in \mathbb{H}^{2,2}_{K+2}$ . Let us denote

$$\underline{\Delta R} = -\partial_t \underline{\Delta P} - \widetilde{\mathcal{L}} \underline{\Delta P}.$$

Then, we can check that if  $\tilde{x} > \overline{X}_P$ ,

$$\begin{split} \underline{\Delta R}_{t}(\widetilde{x}) &= \frac{\epsilon}{1 + \|r\|_{\infty}} \left\{ e^{t-T} \phi(t, \widetilde{x} - \overline{X}_{P}; 0, 1) - \left(1 - e^{t-T}\right) \left[ -r_{t} e^{\widetilde{x} - \overline{X}_{P}} \right. \\ &+ \beta(\widetilde{x} - \overline{X}_{P}) + \left(r_{t} + \frac{|\theta_{t}|^{2}}{2}\right) \right] \right\} \\ &\leq \frac{\epsilon}{1 + \|r\|_{\infty}} \left( e^{\widetilde{x} - \overline{X}_{P}} + |r_{t}| e^{\widetilde{x} - \overline{X}_{P}} \right) \\ &\leq \epsilon e^{\widetilde{x} - \overline{X}_{P}} \\ &\leq \epsilon e^{\widetilde{x}} \\ &\leq \Delta R_{t}(\widetilde{x}). \end{split}$$

If  $\tilde{x} < \overline{X}_P$ , then by (4.8) we have

$$\begin{split} \underline{\Delta R}_{t}(\widetilde{x}) &\leq -Ce^{2M(T-t)} \left\{ e^{(K+1)|\widetilde{x}-\overline{X}_{P}|} \left[ M - \frac{(K+1)^{2}|\theta_{t}|^{2}}{2} \right. \\ &\left. - \left( \beta + |r_{t}| + \frac{|\theta_{t}|^{2}}{2} \right) (K+1) \right] - 2M \right\} \\ &\leq -Ce^{2M(T-t)} \left[ 2e^{(K+1)|\widetilde{x}-\overline{X}_{P}|} - 2M \right], \end{split}$$

provided *M* is large enough. Fix *M* and choose  $\overline{X}_P > 2\log(2MC) + 2 + 2(\log \epsilon)^-$ . Then, we derive that

$$\begin{split} \underline{\Delta R}_{t}(\widetilde{x}) &\leq 2eMC \leq \epsilon e^{\overline{X}_{P}/2} \\ &\leq \epsilon e^{\widetilde{x}} \\ &< \Delta R_{t}(\widetilde{x}), \qquad \text{if } \frac{\overline{X}_{P}}{2} \leq \widetilde{x} < \overline{X}_{P}, \ T - \frac{1}{2M} \leq t \leq T; \\ \underline{\Delta R}_{t}(\widetilde{x}) &\leq -C \left[ e^{(K+1)|\widetilde{x} - \overline{X}_{P}|} + e^{\overline{X}_{P}/2} - 2M \right] \leq -Ce^{(K+1)|\widetilde{x}|} \\ &< \Delta R_{t}(\widetilde{x}), \qquad \text{if } \widetilde{x} < \frac{\overline{X}_{P}}{2}. \end{split}$$

Here we have assumed that C > 1. Until now, we have proved that  $(\underline{\Delta P}, \underline{\Delta Q})$  with  $\underline{\Delta Q} = 0$  is the strong solution to the following BSPDE,

$$\begin{cases} d\underline{\Delta P}_t = -\left(\widetilde{\mathcal{L}}\underline{\Delta P}_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \underline{\Delta Q}_{i,t} + \underline{\Delta R}_t\right) dt + \sum_{i=1}^{N_2} \underline{\Delta Q}_{i,t} dB_{i,t};\\ \underline{\Delta P}_T(\widetilde{x}) \le 0 = \Delta P_T(\widetilde{x}), \end{cases}$$

and  $\underline{\Delta R}_t < \Delta R_t$  if  $T - 1/(2M) \le t \le T$ . Since the strong solution ( $\Delta P, \Delta Q, \Delta K^+$ ) to (5.2) can be written as a strong solution to the following BSPDE,

$$\begin{cases} d\Delta P_t = -\left(\widetilde{\mathcal{L}}\Delta P_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \Delta Q_{i,t} + \Delta R_t + \Delta K_t^+\right) dt + \sum_{i=1}^{N_2} \Delta Q_{i,t} dB_{i,t} \\ \Delta P_T(\widetilde{x}) = 0, \end{cases}$$

where  $\Delta K^+ \ge 0$ . From Lemma 7, we deduce that  $\Delta P_t \ge \underline{\Delta P}_t$  if  $T - 1/(2M) \le t \le T$ . Since  $\underline{\Delta P}_t(\widetilde{x}) > 0$  for any  $\widetilde{x} > \overline{X}_P$  and t < T, we have proved that  $\Delta P_t(\widetilde{x}) > 0$  for any  $\widetilde{x} > \overline{X}_P$  and  $T - 1/(2M) \le t < T$ . Repeating the same argument in the intervals  $[T - (i+1)/(2M), T - i/(2M)], i = 1, 2, \cdots$ , we can deduce that  $\Delta P_t(\widetilde{x}) > 0$  for any  $\widetilde{x} > \overline{X}_P$  and  $0 \le t < T$ .

We will now prove that there exists  $\underline{X}_P$  such that  $\Delta P_t(\widetilde{x}) = 0$  if  $\widetilde{x} \leq \underline{X}_P$ . In order to show this property, we construct a function  $\overline{\Delta P}$  such that  $\overline{\Delta P} \geq \Delta P$  and  $\overline{\Delta P}_t(\widetilde{x}) = 0$  if  $\widetilde{x} \leq \underline{X}_P$ .

From Lemma 9, there exists a negative constant  $\widehat{X}_P$  such that  $\Delta R_t(\widetilde{x}) < -e^{\widetilde{x}}$  for any  $\widetilde{x} \leq \widehat{X}_P$ . Construct  $\overline{\Delta P}$  as the following,

$$\overline{\Delta P}_{t}(\widetilde{x}) = \frac{1}{2M} e^{2\underline{X}_{P}} \phi(t, \widetilde{x} - \underline{X}_{P}; M, K+1) I_{\{\widetilde{x} > \underline{X}_{P}\}} \ge 0,$$

where  $\phi$  is defined in (4.7), *K* is the constant in (5.4), and M > 0,  $\underline{X}_P < 3\widehat{X}_P$  will be determined later. As in the above, we denote  $\overline{\Delta R} \triangleq -\partial_t \overline{\Delta P} - \widetilde{\mathcal{L}} \overline{\Delta P}$ . Then we can show that  $\overline{\Delta P} \in \mathbb{H}^{2,2}_{K+2}$  and

$$\overline{\Delta R}_t(\widetilde{x}) = 0 > \Delta R_t(\widetilde{x}), \quad \forall \quad \widetilde{x} < \underline{X}_P.$$

Moreover, by (4.8), we can check that if  $\tilde{x} > \underline{X}_{P}$ ,

$$\begin{split} \overline{\Delta R}_t(\widetilde{x}) &\geq \frac{e^{2M(T-t)+2\underline{X}_P}}{2M} \left\{ e^{(K+1)|\widetilde{x}-\underline{X}_P|} \left[ M - \frac{(K+1)^2 |\theta_t|^2}{2} - \left(\beta + |r_t| + \frac{|\theta_t|^2}{2}\right)(K+1) \right] - 2M \right\} \\ &\geq \frac{e^{2M(T-t)+2\underline{X}_P}}{2M} \left[ 2e^{(K+1)|\widetilde{x}-\underline{X}_P|} - 2M \right], \end{split}$$

provided *M* is large enough. Fix *M* and choose  $\underline{X}_P < -\log(2MC) - 2MT$ , where *C* is the constant in (5.4). Then, we derive that if  $\underline{X}_P < \tilde{x} < \underline{X}_P/3$ ,

$$\overline{\Delta R}_t(\widetilde{x}) \ge -e^{2M(T-t)+2\underline{X}_P} \ge -e^{\underline{X}_P} \ge -e^{\widetilde{x}} \ge \Delta R_t(\widetilde{x}).$$

Moreover, if  $\tilde{x} > \underline{X}_P/3$ , we have the following estimate,

$$\overline{\Delta R}_{t}(\widetilde{x}) \geq \frac{e^{2M(T-t)+2\underline{X}_{P}}}{2M} \left[ e^{(K+1)(|\widetilde{x}-\underline{X}_{P}|-|\widetilde{x}|)} e^{(K+1)|\widetilde{x}|} \right]$$

Springer

$$+ \left(e^{(K+1)|\widetilde{x}-\underline{X}_{P}|} - 2M\right) \left]$$

$$\geq e^{\frac{2\underline{X}_{P}}{2\underline{M}}} e^{-(K+1)\underline{X}_{P}/3} e^{(K+1)|\widetilde{x}|}$$

$$= \frac{1}{2M} e^{-(K-5)\underline{X}_{P}/3} e^{(K+1)|\widetilde{x}|}$$

$$\geq C e^{(K+1)|\widetilde{x}|} \geq \Delta R_{t}(\widetilde{x}).$$

Here we have supposed that C, M > 1 and K > 8. Until now, we have proved that  $(\overline{\Delta P}, \overline{\Delta Q}, 0)$  with  $\overline{\Delta Q} = 0$  is the strong solution to the following BSPDVI,

$$\begin{cases} d\overline{\Delta P}_{t} = -\left(\widetilde{\mathcal{L}}\overline{\Delta P}_{t} + \sum_{i=1}^{N_{2}}\widetilde{\mathcal{M}}_{i}\overline{\Delta Q}_{i,t} + \overline{\Delta R}_{t}\right) dt \\ + \sum_{i=1}^{N_{2}}\overline{\Delta Q}_{i,t} dB_{i,t} & \text{if } \overline{\Delta P} > 0; \\ d\overline{\Delta P}_{t} \leq -\left(\widetilde{\mathcal{L}}\overline{\Delta P}_{t} + \sum_{i=1}^{N_{2}}\widetilde{\mathcal{M}}_{i}\overline{\Delta Q}_{i,t} + \overline{\Delta R}_{t}\right) dt \\ + \sum_{i=1}^{N_{2}}\overline{\Delta Q}_{i,t} dB_{i,t} & \text{if } \overline{\Delta P} = 0; \\ \overline{\Delta P}_{T}(\widetilde{x}) \geq 0 \geq P_{T}(\widetilde{x}), \end{cases}$$

and  $\overline{\Delta R} \ge \Delta R$ . Applying Lemma 5, we deduce that  $\Delta P \le \overline{\Delta P}$ . Since  $\overline{\Delta P}_t(\widetilde{x}) = 0$  for any  $\widetilde{x} \le \underline{X}_P$ , we have proved that  $\Delta P_t(\widetilde{x}) = 0$  if  $\widetilde{x} \le \underline{X}_P$ .  $\Box$ 

**Theorem 10** If  $\gamma + \rho \leq 1$ , then we have  $\Delta P$  is increasing with respect to  $\widetilde{x}$ .

Proof Denote

$$\Delta P_t^{\lambda}(\widetilde{x}) = \Delta P_t(\widetilde{x} + \lambda), \quad \Delta Q_t^{\lambda}(\widetilde{x}) = \Delta Q_t(\widetilde{x} + \lambda), \quad \Delta R_t^{\lambda}(\widetilde{x}) = \Delta R_t(\widetilde{x} + \lambda),$$

where  $\lambda$  is an arbitrary positive number such that  $\lambda < 1$ . Then  $\Delta P^{\lambda}$ ,  $\Delta Q^{\lambda}$  satisfies the following BSPDVI

$$\begin{cases} d\Delta P_t^{\lambda} = -\left(\widetilde{\mathcal{L}}\Delta P_t^{\lambda} + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \Delta Q_{i,t}^{\lambda} + \Delta R_t^{\lambda}\right) dt \\ + \sum_{i=1}^{N_2} \Delta Q_{i,t}^{\lambda} dB_{i,t} & \text{if } \Delta P^{\lambda} > 0; \\ d\Delta P_t^{\lambda} \le -\left(\widetilde{\mathcal{L}}\Delta P_t^{\lambda} + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i \Delta Q_{i,t}^{\lambda} + \Delta R_t^{\lambda}\right) dt \\ + \sum_{i=1}^{N_2} \Delta Q_{i,t}^{\lambda} dB_{i,t} & \text{if } \Delta P^{\lambda} = 0; \\ \Delta P_T^{\lambda}(\widetilde{x}) = 0. \end{cases}$$

Since  $\Delta R$  is increasing with respect to  $\tilde{x}$  by Lemma 11, we have  $\Delta R^{\lambda} \geq \Delta R$ . Recalling Lemma 5, we deduce that  $\Delta P^{\lambda} \geq \Delta P$ . Hence, we have shown that  $\Delta P$  is increasing with respect to  $\tilde{x}$ .

From Theorem 9, Theorem 10 and (4.1), (5.1), we have the following properties of the dual value  $\hat{V}$ .

**Theorem 11** (1) There exist two positive constants  $\overline{X} = e^{\overline{X}_P}$ ,  $\underline{X} = e^{\underline{X}_P}$  such that  $\widehat{V}_t(x) > \underline{\widehat{V}}_t(x)$  if  $x > \overline{X}$  and t < T, and  $\widehat{V}_t(x) = \underline{\widehat{V}}_t(x)$  if  $x \leq \underline{X}$ . (2) If  $\gamma + \rho \leq 1$ , then  $\widehat{V} - \underline{\widehat{V}}$  is increasing with respect to x.

Define the optimal retirement boundary in t - x coordinate system as follows:

$$\mathcal{R}_t^x \triangleq \inf\{x > 0 : \widehat{V}_t(x) > \underline{\widehat{V}}_t(x)\}, \ t \in [0, T).$$

Define the working region and the retirement region as follows:

$$\mathbf{WR}^{x} \triangleq \{(\omega, t, x) : \widehat{V}_{t}(\omega, x) > \underline{\widehat{V}}_{t}(\omega, x)\},\\ \mathbf{RR}^{x} \triangleq \{(\omega, t, x) : \widehat{V}_{t}(\omega, x) = \underline{\widehat{V}}_{t}(\omega, x)\}.$$

From Theorem 11, and we have the following properties of the optimal retirement boundary in t - x coordinate system as the follows.

**Theorem 12** (1) There exist two positive constants  $\overline{X}$ ,  $\underline{X}$  such that  $\underline{X} \leq \mathcal{R}^x \leq \overline{X}$ . (2) If  $\gamma + \rho \leq 1$ , then  $\mathbf{WR}^x = \{(\omega, t, x) : x > \mathcal{R}^x_t(\omega), t \in [0, T)\}$  and  $\mathbf{RR}^x = \{(\omega, t, x) : 0 < x \leq \mathcal{R}^x_t(\omega)\}$ .

If the initial  $(\omega, t, x) \in \mathbf{WR}^x$ , then  $\widehat{V}(\omega, t, x) > \widehat{V}(\omega, t, x)$  a.s. in  $\Omega$ . The definition of  $\tau^*$  in Theorem 1 implies that  $\tau^*(\omega) > t$  a.s. in  $\Omega$ , i.e., the agent chooses to work. As time passes, s > t, before the trajectory of the dual variable process  $X_s^*$  first hits the optimal retirement boundary  $\mathcal{R}_s^x$ ,  $(\omega, s, X_s^*(\omega)) \in \mathbf{WR}^x$  and  $\tau^*(\omega) > s$ , i.e., the agent keeps working. If  $X_s^*$  hits  $\mathcal{R}_s^x$ , then  $(\omega, s, X_s^*(\omega)) \in \mathbf{RR}^x$ ,  $\tau^*(\omega) = s$  and the agent chooses to retire. If the initial  $(\omega, t, x) \in \mathbf{RR}^x$ , however, then  $\widehat{V}(\omega, t, x) = \widehat{V}(\omega, t, x)$ a.s. in  $\Omega$ . The definition of  $\tau^*$  in Theorem 1 implies that  $\tau^*(\omega) = t$  a.s. in  $\Omega$ , i.e., the agent chooses to retire.

Theorem 12(a) provides constant bounds, namely, an upper bound  $\overline{X}$  and a lower bound  $\underline{X}$  for the boundary. Theorem 12(b) provides a characterization of the working region and the retirement region in terms of the optimal retirement boundary. In particular, it says that under the condition that  $\gamma + \rho \leq 1$  the agent retires at the first time when the marginal utility of wealth  $X_s$  hits the optimal retirement boundary. That is, under the condition that the coefficient of relative risk aversion is less than or equal to the reciprocal of the elasticity of substitution between consumption and leisure ( $\gamma + \rho \leq 1$ ), the agent retires when the agent's marginal utility becomes sufficiently small to hit the boundary.

Next, we will come back to study the optimal retirement threshold in the original coordinate system (t, y), where y denotes the wealth of the agent. For this purpose, we redefine the working domain, the retirement domain and define the optimal retirement threshold in the (t, y)-coordinate system.

Recalling Theorem 1, we know that  $x_t^*(y) = \mathcal{J}_{\widehat{V},t}(-y - \mathcal{Y}_t)$  is continuous and strictly decreasing with respect to y a.s. in  $\Omega \times [0, T]$ , and is a bijection from  $(-\mathcal{Y}_t(\omega), +\infty)$  to  $(0, +\infty)$  a.s. in  $\Omega \times [0, T]$ . So  $x_t^*(\omega, \cdot)$  has an inverse function  $y_t^*(\omega, \cdot)$ , which is continuous, strictly decreasing and maps  $(0, +\infty)$  to  $(-\mathcal{Y}_t(\omega), +\infty)$  a.s. in  $\Omega \times [0, T]$ .

Let us define the optimal retirement threshold as  $\mathcal{R}_t^y \triangleq y_t^*(\mathcal{R}_t^x)$ , and the working domain and the retirement domain in t - y as the following,

$$\mathbf{R}\mathbf{R}^{y} \triangleq \{(\omega, t, y) : (\omega, t, x_{t}^{*}(\omega, y)) \in \mathbf{R}\mathbf{R}^{x}\},\$$
$$\mathbf{W}\mathbf{R}^{y} \triangleq \{(\omega, t, y) : (\omega, t, x_{t}^{*}(\omega, y)) \in \mathbf{W}\mathbf{R}^{x}\}.$$

Then we state the properties of the optimal retirement boundary  $\mathcal{R}^{y}$  as the follows.

**Theorem 13** (1) The optimal retirement threshold is given by  $y = \mathcal{R}_t^y = -\partial_x \underline{\widetilde{V}}_t(\mathcal{R}_t^x)$ a.e. in  $\Omega \times [0, T]$ . And there exist two stochastic processes  $\underline{\mathcal{Y}}, \overline{\mathcal{Y}}$  such that  $0 < \underline{\mathcal{Y}} \leq \mathcal{R}^y \leq \overline{\mathcal{Y}}$ , where  $\underline{\mathcal{Y}}_t = -\partial_x \underline{\widetilde{V}}_t(\overline{X})$  and  $\overline{\mathcal{Y}}_t = -\partial_x \underline{\widetilde{V}}_t(\underline{X})$ .

(2) If  $\gamma + \rho \leq 1$ , then  $\mathbf{WR}^{y} = \{(\omega, t, y) : -\mathcal{Y}_{t} < y < \mathcal{R}_{t}^{y}(\omega), t \in [0, T)\}$  and  $\mathbf{RR}^{y} = \{(\omega, t, y) : y \geq \mathcal{R}_{t}^{y}(\omega)\}.$ 

**Proof** It is sufficient to prove that  $\mathcal{R}_t^y = -\partial_x \underline{\widetilde{V}}_t(\mathcal{R}_t^x)$ , and other statements come from Theorem 12.

Recalling  $x_t^*(y) = \mathcal{J}_{\widehat{V},t}(-y - \mathcal{Y}_t)$ , we have  $\partial_x \widehat{V}_t(x_t^*(y)) = -y - \mathcal{Y}_t$ . Taking  $y = y_t^*(x)$ , we can show by computation

$$\partial_x \widehat{V}_t(x) = \partial_x \widehat{V}_t(x_t^*(y_t^*(x))) = -y_t^*(x) - \mathcal{Y}_t \text{ for any } x > 0 \text{ a.e. in } \Omega \times [0, T].$$

Hence, we deduce that  $y_t^*(x) = -\mathcal{Y}_t - \partial_x \widehat{V}_t(x)$  for any x > 0 a.e. in  $\Omega \times [0, T]$ . So we have

$$\mathcal{R}_{t}^{y} = y_{t}^{*}(\mathcal{R}_{t}^{x}) = -\mathcal{Y}_{t} - \partial_{x}\widehat{V}_{t}(\mathcal{R}_{t}^{x}) = -\mathcal{Y}_{t} - \partial_{x}\underline{\widehat{V}}_{t}(\mathcal{R}_{t}^{x})$$
$$= -\mathcal{Y}_{t} - \partial_{x}\left[\underline{\widetilde{V}}_{t}(x) - x\mathcal{Y}_{t}\right]_{x = \mathcal{R}_{t}^{x}} = -\partial_{x}\underline{\widetilde{V}}_{t}(\mathcal{R}_{t}^{x}) \text{ a.e. in } \Omega \times [0, T],$$

where we have used the fact that  $\partial_x \widehat{V}_t(\mathcal{R}_t^x) = \partial_x \underline{\widehat{V}}(\mathcal{R}_t^x)$  in the third equality, which can be derived from the continuity of  $\partial_x \widehat{V}$ ,  $\partial_x \underline{\widehat{V}}$  with respect to x and  $\widehat{V}_t(x) = \underline{\widehat{V}}_t(x)$  for any  $x \leq \mathcal{R}_t^x$ . And we have used the fact  $\underline{\widehat{V}} = \underline{\widetilde{V}} - x \mathcal{Y}_t$  in the fourth equality.  $\Box$ 

If the trajectory of the wealth process  $Y_s^*$  stays in  $\mathbf{WR}^y$ , then the trajectory of the dual process  $X_s^*$  stays in  $\mathbf{WR}^x$  and the agent chooses to work. If, however, the trajectory of the wealth process  $Y_s^*$  reaches  $\mathbf{RR}^y$ , then the trajectory of the dual variable process  $X_s^*$  reaches  $\mathbf{RR}^x$  and the agent chooses to retire. Theorem 13(a) provides bounds for the optimal retirement threshold. In contrast to the bounds for the optimal retirement boundary in the dual space, the bounds are stochastic. Theorem 13(b) provides a characterization of the working region and the retirement region in terms of the optimal retirement threshold. In particular, it says that under the condition that  $\gamma + \rho \leq 1$  the agent retires at the first time when wealth  $X_s$  hits the level of the optimal retirement

threshold. That is, under the condition that the coefficient of relative risk aversion is less than or equal to the reciprocal of the elasticity of substitution between consumption and leisure ( $\gamma + \rho \leq 1$ ), the agent retires when the agent's wealth becomes sufficiently large to reach the threshold level.

The optimal retirement threshold is such that the agent retires as soon as the wealth process reaches the threshold.

## 6 Conclusion

We have studied an optimal retirement and consumption/portfolio selection problem of an agent in a non-Markovian market environment. We have derived a BSPDVI by applying successive transformations and provided a verification theorem that the solution to the original optimization problem can be obtained by a strong solution to the BSPDVI satisfying certain regularity conditions. We have also shown that there exists a unique strong solution to the BSPDVI satisfying the conditions.

Since the theory of BSPDE is not mature, we have relied on a combination of a probabilistic approach and the theory. Derivation of the result by using solely the theory of BSPDE seems an interesting topic to pursue. We have not been able to a full concrete characterization of the early retirement boundary, as is done in Yang and Koo [30]. A full characterization of the boundary is left as future research.

## A A Few Existing Results About BSPDVI

For convenience of the reader we state the generalized Itô-Kunita-Wentzell's formula, and results about BSPDE (refer to [8,31]) and BSPDVI (refer to [21]) which were used in the paper.

We provide the generalized Itô–Kunita–Wentzell's formula (refer to Theorem 3.1 in [31]) in the following<sup>9</sup>,

**Lemma 3** Suppose that the random function  $v : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$  satisfies the following properties: v(x) is a continuous semimartingale and takes the following form

$$v_s(x) = v_t(x) + \int_t^s f_u(x) \, du + \sum_{i=1}^{N_2} \int_t^s w_{i,u}(x) \, dB_{i,u}$$
  
a.e.  $x \in \mathbb{R}$  for every  $s \in [t, T]$  a.s. in  $\Omega$ .

And  $v \in \mathbb{H}^{2,2}(D)$ ,  $w \in \mathbb{H}^{1,2}(D)$ ,  $f \in \mathbb{H}^{0,2}(D)$  for any compact subset D of  $\mathbb{R}^+$ . Let X be governed by (3.9), and Assumptions 2 and 3 be satisfied. Then we have

<sup>&</sup>lt;sup>9</sup> Though  $\beta = 0$  in Theorem 3.1 in [31], it is clear that the corresponding result still holds in the case of  $\beta \neq 0$ . Moreover,  $v \in \mathbb{H}^{2,2}$ ,  $w \in \mathbb{H}^{1,2}$ , and  $f \in \mathbb{H}^{0,2}$  are assumed in Theorem 3.1 in [31]. We relax the assumption, as we can construct a sequence of intervals  $\{[1/n, n] : n \in \mathbb{Z}^+\}$  which converges to  $\mathbb{R}^+$ .

$$e^{-\beta(s-t)}v_{s}(X_{s}) = v_{t}(X_{t}) + \int_{t}^{s} e^{-\beta(u-t)} \left( \mathcal{L}v_{u} + \sum_{i=1}^{N_{2}} \mathcal{M}_{i}w_{i,u} + f_{u} \right) (X_{u}) du$$
  
+ 
$$\int_{t}^{s} e^{-\beta(u-t)} \left[ \sum_{i=1}^{N_{2}} (w_{i,u} + \mathcal{M}_{i}v_{u}) (X_{u}) dB_{i,u} - \partial_{x}v_{u}(X_{u})X_{u} \left( \theta_{u}^{1} \right)^{\top} dW_{u} \right],$$

where the differential operator  $\mathcal{L}$  and  $\mathcal{M}$  are defined in (3.11).

**Remark 6** From the above integral representation of  $e^{-\beta(\cdot-t)}v_{\cdot}(X_{\cdot})$ , we know that the stochastic process  $v_{\cdot}(X_{\cdot})$  has a continuous version.

We will now state results about a BSPDVI, which takes the following form

$$\begin{cases} dv_t = -\left(\widetilde{\mathcal{L}}v_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i w_{i,t} + f_t\right) dt + \sum_{i=1}^{N_2} w_{i,t} dB_{i,t} & \text{if } v_t > \underline{v}_t; \\ dv_t \le -\left(\widetilde{\mathcal{L}}v_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i w_{i,t} + f_t\right) dt + \sum_{i=1}^{N_2} w_{i,t} dB_{i,t} & \text{if } v_t = \underline{v}_t; \\ v_T(x) = \varphi(x), \end{cases}$$
(A.1)

where the differential operator  $\widetilde{\mathcal{L}}$  and  $\widetilde{\mathcal{M}}$  are defined in (4.3).

First, we state a result about the existence and uniqueness of the strong solution to the BSPDVI (see Theorem 5.4 in [21]).

**Lemma 4** Let Assumptions 2 and 3 be satisfied,  $f \in \mathbb{H}^{0,2}_{\lambda}$ ,  $\varphi \in \mathbb{L}^{1,2}_{\lambda}$ ,  $\underline{v} \in \mathbb{H}^{2,2}_{\lambda}$ with some nonnegative number  $\lambda$ . Moreover, the lower obstacle  $\underline{v}$  is a continuous semimartingale of the following form

$$d\underline{v}_t = -\underline{g}_t dt + \sum_{i=1}^{N_2} \underline{w}_{i,t} dB_{i,t}$$

with some  $(\underline{w}, \underline{g}) \in \mathbb{H}^{1,2}_{\lambda} \times \mathbb{H}^{0,2}_{\lambda}$ , and  $\underline{v}_T \leq \varphi$ . Then BSPDVI (A.1) has a unique strong solution  $(v, w, k^+) \in \mathbb{H}^{2,2}_{\lambda} \times \mathbb{H}^{1,2}_{\lambda} \times \mathbb{H}^{0,2}_{\lambda}$ . Moreover,  $v \in \mathbb{S}^{1,2}_{\lambda}$ .

Next, we provide a comparison theorem for BSPDVI (refer to Theorem 5.2 in [21]).

**Lemma 5** Let the assumptions in Lemma 4 be satisfied. Let  $(v_i, w^i, k_i^+)$  be the strong solution to BSPDVI (A.1) associated with  $(f_i, \varphi_i, \underline{v}_i)$  for i = 1, 2. If  $f_1 \ge f_2, \varphi_1 \ge \varphi_2$ , and  $\underline{v}_1 \ge \underline{v}_2$ , then  $v_1 \ge v_2$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}$ .

Consider the following BSPDE,

$$\begin{cases} dv_t = -\left(\widetilde{\mathcal{L}}v_t + \sum_{i=1}^{N_2} \widetilde{\mathcal{M}}_i w_{i,t} + f_t\right) dt + \sum_{i=1}^{N_2} w_{i,t} dB_{i,t};\\ v_T(x) = \varphi(x). \end{cases}$$
(A.2)

In view of the results for BSPDE (refer to Lemmas 2.2 and 5.1 in [31], or Theorem 5.3 and Corollary 3.4 in [8]), we can conclude the following results for  $BSPDE^{10}$ .

**Lemma 6** Let Assumptions 2 and 3 be satisfied,  $f \in \mathbb{H}^{0,2}_{\lambda}$ ,  $\varphi \in \mathbb{L}^{1,2}_{\lambda}$  with some nonnegative number  $\lambda$ . Then BSPDE (A.2) has a unique strong solution  $(v, w) \in \mathbb{H}^{2,2}_{\lambda} \times \mathbb{H}^{1,2}_{\lambda}$ . Moreover,  $v \in \mathbb{S}^{1,2}_{\lambda}$ .

**Lemma 7** Let the assumptions in Lemma 6 be satisfied. Assume that  $(v_i, w^i)$  be the strong solution to BSPDE (A.2) associated with  $(f_i, \varphi_i)$  for i = 1, 2. If  $f_1 \ge f_2$  and  $\varphi_1 \ge \varphi_2$ , then  $v_1 \ge v_2$  a.e. in  $\Omega \times [0, T] \times \mathbb{R}$ .

Finally, we give the result about the relationship between the strong solution  $\widehat{P}$  of BSPDVI (4.4) and the value of the corresponding optimal stopping problem (refer to Theorem 5.4 in [21])<sup>11</sup>.

**Lemma 8** Let Assumptions 2 and 3 be satisfied. Suppose that  $\underline{\widetilde{P}}$  and  $\widehat{P}$  are the strong solutions to BSPDE (4.2) and BSPDVI (4.4), respectively. The terminal value and lower obstacle in BSPDVI (4.4),  $\underline{\widehat{P}}_{t}(\widetilde{x}) = \underline{\widetilde{P}}_{t}(\widetilde{x}) - e^{\widetilde{x}}\mathcal{Y}_{t}$ . Let  $\widetilde{R}_{1}, \widehat{R}_{1} \in \mathbb{H}^{0,2}_{\lambda}, \widetilde{R}_{2} \in \mathbb{L}^{1,2}_{\lambda}$  be satisfied with some nonnegative number  $\lambda$ . Moreover, suppose that  $\underline{\widehat{P}}(\widetilde{X}^{t,\widetilde{x}}), \ \widehat{P}(\widetilde{X}^{t,\widetilde{x}}) \in S^{2}, \ \widehat{R}_{1}(\widetilde{X}^{t,\widetilde{x}}) \in \mathcal{L}^{2}$ , where  $\widetilde{X}^{t,\widetilde{x}} = \log X^{t,x}$ <sup>12</sup>. Then

$$\underline{\widetilde{P}}_{t}(\widetilde{x}) = \mathbb{E}\left[\int_{t}^{T} e^{-\beta(s-t)} \widetilde{R}_{1}(\widetilde{X}_{s}^{t,\widetilde{x}}) ds + e^{-\beta(T-t)} \widetilde{R}_{2}(\widetilde{X}_{T}^{t,\widetilde{x}}) \left| \mathcal{F}_{t} \right],$$

and  $\widehat{P}_t(\widetilde{x})$  is the value of the following optimal stopping problem,

$$\widehat{P}_{t}(\widetilde{x}) = \underset{\tau \in \mathcal{U}_{t,T}}{\operatorname{ess.sup}} \mathbb{E}\left[\int_{t}^{\tau \wedge T} e^{-\beta(s-t)} \widehat{R}_{1,s}(\widetilde{X}_{s}^{t,\widetilde{x}}) ds + e^{-\beta(\tau-t)} \underline{\widehat{P}}_{\tau \wedge T}(\widetilde{X}_{\tau \wedge T}^{t,\widetilde{x}}) \left| \mathcal{F}_{t} \right.\right]$$

and its optimal stopping time  $\tau^*$  can be described as

$$\tau^* \triangleq \inf \left\{ s \in [t, T] : \widehat{P}_s(\widetilde{X}_s^{t, \widetilde{x}}) = \underline{\widehat{P}}_s(\widetilde{X}_s^{t, \widetilde{x}}) \right\} \wedge T.$$

<sup>&</sup>lt;sup>10</sup> In the Lemmas 2.2 and Lemma 5.1 in [31], the parameter  $\lambda$  in the random field space  $\mathbb{H}_{\lambda}^{i,2}$ , i = 0, 1, 2 is supposed to be zero. But we can utilize the transformation in the proof of Theorem 4.9 in [21] to relax the condition. Moreover, we can obtain the results in Lemmas 6 and 7 by another method as follows. Applying the method in [21], we can construct a  $\underline{v}$  such that  $\underline{v} < v$  under the conditions in Lemma 4. So BSPDVI (A.1) is equivalent to BSPDE (A.2), and Lemmas 6 and 7 follow from Lemmas 4 and 5, respectively.

<sup>&</sup>lt;sup>11</sup> Though we have proved the relationship between  $\hat{V}$  defined in (3.14) and the strong solution  $\hat{v}$  to BSPDVI (3.16) in Theorem 1, the assumptions in Theorem 1 is too strong, we need the relationship under more weaker assumptions.

<sup>&</sup>lt;sup>12</sup> In [21],  $\widehat{R}_1(x)$  is supposed to be of polynomial growth with respect to x, which guarantees that  $\widehat{R}_1(\widetilde{X}^{t,\widetilde{x}}) \in \mathcal{L}^2$ . Now, we relax this condition to  $\widehat{R}_1(\widetilde{X}^{t,\widetilde{x}}) \in \mathcal{L}^2$ .

## **B** Properties of $\Delta R$

In this section, we deduce properties about  $\Delta R$  defined in (5.3), which is important to derive the properties of the optimal retirement boundary.

**Lemma 9** ess.sup{ $\Delta R_t(\widetilde{x}) : (\omega, t) \in \Omega \times [0, T]$ } <  $-e^{\widetilde{x}}$  provided  $\widetilde{x}$  is small enough.

**Proof** We prove the conclusion by considering the property of  $\Delta R_t(\tilde{x})$  as  $\tilde{x} \to -\infty$  in six different cases.

1. Consider the case of  $0 < \rho < 1$  and  $0 < \gamma \neq 1$ . As  $\tilde{x} \to -\infty$ , i.e.,  $x \to 0^+$ , then  $c^* = \mathcal{J}_{U_1}(x; l) \to +\infty$ , and

$$\begin{split} x &= \partial_{c} U_{1}(c^{*}, l) \\ &= \alpha (c^{*})^{\rho-1} \left[ \alpha (c^{*})^{\rho} + (1-\alpha) l^{\rho} \right]^{\frac{1-\gamma-\rho}{\rho}} \\ &= \alpha^{\frac{1-\gamma}{\rho}} (c^{*})^{-\gamma} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{l^{\rho}}{(c^{*})^{\rho}} \right]^{\frac{1-\gamma-\rho}{\rho}} \\ &= \alpha^{\frac{1-\gamma}{\rho}} (c^{*})^{-\gamma} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{1-\gamma-\rho}{\rho} l^{\rho} (c^{*})^{-\rho} + o\left( (c^{*})^{-\rho} \right) \right]. \end{split}$$

So, we deduce that as  $x \to 0^+$ ,

$$\begin{aligned} \mathcal{J}_{U_{1}}(x;l) &= c^{*} \\ &= \alpha^{\frac{1-\gamma}{\rho\gamma}} x^{\frac{-1}{\gamma}} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{1-\gamma-\rho}{\rho} l^{\rho} (c^{*})^{-\rho} + o\left((c^{*})^{-\rho}\right) \right]^{\frac{1}{\gamma}} \\ &= \alpha^{\frac{1-\gamma}{\rho\gamma}} x^{\frac{-1}{\gamma}} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{1-\gamma-\rho}{\rho\gamma} l^{\rho} (c^{*})^{-\rho} + o\left((c^{*})^{-\rho}\right) \right] \\ &= \alpha^{\frac{1-\gamma}{\rho\gamma}} x^{\frac{-1}{\gamma}} \left[ 1 + \frac{1-\gamma-\rho}{\rho\gamma} (1-\alpha) l^{\rho} \alpha^{\frac{-1}{\gamma}} x^{\frac{\rho}{\gamma}} + o\left(x^{\frac{\rho}{\gamma}}\right) \right]. \end{aligned} (B.1)$$

And by computation we can show that as  $x \to 0^+$ ,

$$U_{1}(c^{*},l) - c^{*}x = \frac{\alpha^{\frac{1-\gamma}{\rho}}(c^{*})^{1-\gamma}}{1-\gamma} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{l^{\rho}}{(c^{*})^{\rho}} \right]^{\frac{1-\gamma}{\rho}} - c^{*}x$$
$$= \frac{\alpha^{\frac{1-\gamma}{\rho}}(c^{*})^{1-\gamma}}{1-\gamma} \left[ 1 + \frac{1-\alpha}{\alpha} \frac{1-\gamma}{\rho} l^{\rho}(c^{*})^{-\rho} + o\left((c^{*})^{-\rho}\right) \right]$$
$$-c^{*}x$$
$$= \frac{\alpha^{\frac{1-\gamma}{\rho}}}{1-\gamma} \alpha^{\frac{(1-\gamma)^{2}}{\rho\gamma}} x^{\frac{\gamma-1}{\gamma}} \left[ 1 + \frac{1-\gamma-\rho}{\rho\gamma} (1-\alpha)(1-\gamma) l^{\rho} \alpha^{\frac{-1}{\gamma}} x^{\frac{\rho}{\gamma}} + o\left(x^{\frac{\rho}{\gamma}}\right) \right] \left[ 1 + \frac{1-\alpha}{\alpha} \frac{1-\gamma}{\rho} l^{\rho} \alpha^{\frac{\gamma-1}{\gamma}} x^{\frac{\rho}{\gamma}} + o\left(x^{\frac{\rho}{\gamma}}\right) \right]$$

$$-\alpha^{\frac{1-\gamma}{\rho\gamma}}x^{\frac{\gamma-1}{\gamma}}\left[1+\frac{1-\gamma-\rho}{\rho\gamma}(1-\alpha)l^{\rho}\alpha^{\frac{-1}{\gamma}}x^{\frac{\rho}{\gamma}}+o\left(x^{\frac{\rho}{\gamma}}\right)\right]$$
$$=\alpha^{\frac{1-\gamma}{\rho\gamma}}x^{\frac{\gamma-1}{\gamma}}\left[\frac{\gamma}{1-\gamma}+\frac{1-\alpha}{\rho}l^{\rho}\alpha^{\frac{-1}{\gamma}}x^{\frac{\rho}{\gamma}}+o\left(x^{\frac{\rho}{\gamma}}\right)\right].$$

Note that  $\overline{x}_t$  has a positive lower bound. Hence, we have that as  $x \to 0^+$ ,

$$\begin{split} \Delta R_t(\widetilde{x}) &= \widehat{U}_{1,t}(x) - \widetilde{U}_{1,t}(x) + L_t w_t x \\ &= U_1(\mathcal{J}_{U_1}(x;\overline{L}_t),\overline{L}_t) - x(\mathcal{J}_{U_1}(x;\overline{L}_t) + \overline{L}_t w_t) - U_1(\mathcal{J}_{U_1}(x;L_t),L_t) \\ &+ x\mathcal{J}_{U_1}(x;L_t) + L_t w_t x \\ &= \frac{1-\alpha}{\rho} \alpha^{\frac{1-\gamma-\rho}{\rho\gamma}} x^{1-\frac{1-\rho}{\gamma}} \left[ \overline{L}_t^{\ \rho} - L_t^{\ \rho} + o(1) \right] + (L_t - \overline{L}_t) w_t x \\ &= \frac{1-\alpha}{\rho} \alpha^{\frac{1-\gamma-\rho}{\rho\gamma}} x^{1-\frac{1-\rho}{\gamma}} \left[ \overline{L}_t^{\ \rho} - L_t^{\ \rho} + o(1) \right] \\ &< -x \\ &= -e^{\widetilde{x}} \end{split}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough, where we have used Assumption 2.

2. In the case of  $0 < \rho < 1$  and  $\gamma = 1$ , (B.1) still holds. And by computation we can show that as  $x \to 0^+$ ,

$$U_{1}(c^{*}, l) - c^{*}x = \frac{\log \alpha}{\rho} + \log c^{*} + \frac{1}{\rho} \log \left[ 1 + \frac{1 - \alpha}{\alpha} \frac{l^{\rho}}{(c^{*})^{\rho}} \right] - c^{*}x$$
$$= \frac{\log \alpha}{\rho} - \log x - \frac{1 - \alpha}{\alpha} l^{\rho} x^{\rho} + \frac{1 - \alpha}{\rho \alpha} l^{\rho} x^{\rho}$$
$$- \left[ 1 - \frac{1 - \alpha}{\alpha} l^{\rho} x^{\rho} \right] + o(x^{\rho})$$
$$= \frac{\log \alpha}{\rho} - \log x - 1 + \frac{1 - \alpha}{\rho \alpha} l^{\rho} x^{\rho} + o(x^{\rho}),$$

and

$$\Delta R_t(\widetilde{x}) = \frac{1-\alpha}{\rho\alpha} x^{\rho} \left[ \overline{L_t}^{\rho} - L_t^{\rho} + o(1) \right] + (L_t - \overline{L}_t) w_t x < -x = -e^{\widetilde{x}}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough.

3. In the case of  $\rho = 0$  and  $0 < \gamma \neq 1$ , the proof proceeds similarly to the above. Since

$$x = \partial_c U_1(c^*, l) = \alpha(c^*)^{\alpha(1-\gamma)-1} l^{(1-\alpha)(1-\gamma)},$$

we have that

$$\mathcal{J}_{U_1}(x;l) = c^* = \alpha^{\frac{1}{1-\alpha(1-\gamma)}} l^{\frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)}} x^{\frac{1}{\alpha(1-\gamma)-1}}.$$
 (B.2)

And by computation we can show that

$$U_1(c^*, l) - c^* x = \left[\frac{1}{1 - \gamma} \alpha^{\frac{\alpha(1 - \gamma)}{1 - \alpha(1 - \gamma)}} - \alpha^{\frac{\alpha}{1 - \alpha(1 - \gamma)}}\right] l^{\frac{(1 - \alpha)(1 - \gamma)}{1 - \alpha(1 - \gamma)}} x^{\frac{\alpha(1 - \gamma)}{\alpha(1 - \gamma) - 1}}.$$

Note that if  $\gamma > 1$ , then

$$0 < \frac{\alpha(1-\gamma)}{\alpha(1-\gamma)-1} < 1, \quad \frac{1}{1-\gamma} \alpha^{\frac{\alpha(1-\gamma)}{1-\alpha(1-\gamma)}} - \alpha^{\frac{\alpha}{1-\alpha(1-\gamma)}} < 0, \quad \frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)} < 0,$$

and if  $0 < \gamma < 1$ , then

$$\frac{\alpha(1-\gamma)}{\alpha(1-\gamma)-1} < 0, \ \frac{1}{1-\gamma} \left(\alpha^{\frac{\alpha}{1-\alpha(1-\gamma)}}\right)^{1-\gamma} - \alpha^{\frac{\alpha}{1-\alpha(1-\gamma)}} > 0, \ \frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)} > 0.$$

We can conclude that  $U_1(c^*, l) - c^*x$  is strictly increasing with respect to l, and

$$\Delta R_{t}(\widetilde{x}) = \left[\frac{1}{1-\gamma}\alpha^{\frac{\alpha(1-\gamma)}{1-\alpha(1-\gamma)}} - \alpha^{\frac{\alpha}{1-\alpha(1-\gamma)}}\right] \left[\overline{L_{t}}^{\frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)}} - L_{t}^{\frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)}} + o(1)\right]$$

$$\times x^{\frac{\alpha(1-\gamma)}{\alpha(1-\gamma)-1}}$$

$$< -x$$

$$= -e^{\widetilde{x}}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough.

4. In the case of  $\rho = 0$  and  $\gamma = 1$ , (B.2) still holds and similarly to the above we have

$$U_1(c^*, l) - c^* x = \alpha \log \alpha - \alpha \log x + (1 - \alpha) \log l - \alpha,$$
  
$$\Delta R_t(\widetilde{x}) = (1 - \alpha) (\log \overline{L}_t - \log L_t) + (L_t - \overline{L}_t) w_t x < -x = -e^{\widetilde{x}}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough.

5. In the case of  $\rho < 0$  and  $0 < \gamma \neq 1$ , similarly to the above, as  $\tilde{x} \to -\infty$ , i.e.,  $x \to 0^+, c^* = \mathcal{J}_{U_1}(x; l) \to +\infty$ , and

$$\begin{aligned} x &= \partial_c U_1(c^*, l) \\ &= \alpha(c^*)^{\rho-1} [\alpha(c^*)^{\rho} + (1-\alpha)l^{\rho}]^{\frac{1-\gamma-\rho}{\rho}} \\ &= \alpha(c^*)^{\rho-1} [o(1) + (1-\alpha)l^{\rho}]^{\frac{1-\gamma-\rho}{\rho}} \\ &= \alpha(1-\alpha)^{\frac{1-\gamma-\rho}{\rho}} l^{1-\gamma-\rho} (c^*)^{\rho-1} [1+o(1)] \end{aligned}$$

So we deduce that as  $x \to 0^+$ ,

$$\mathcal{J}_{U_1}(x;l) = c^* = \alpha^{\frac{1}{1-\rho}} (1-\alpha)^{\frac{1-\gamma-\rho}{\rho(1-\rho)}} l^{\frac{1-\gamma-\rho}{1-\rho}} x^{\frac{1}{\rho-1}} [1+o(1)].$$
(B.3)

And by computation we can show that as  $x \to 0^+$ ,

$$U_{1}(c^{*}, l) - c^{*}x = \frac{(1-\alpha)^{\frac{1-\gamma}{\rho}}l^{1-\gamma}}{1-\gamma} \left[1+o(1)\right]^{\frac{1-\gamma}{\rho}} \\ -\alpha^{\frac{1}{1-\rho}}(1-\alpha)^{\frac{1-\gamma-\rho}{\rho(1-\rho)}}l^{\frac{1-\gamma-\rho}{1-\rho}}x^{\frac{\rho}{\rho-1}}\left[1+o(1)\right] \\ = \frac{(1-\alpha)^{\frac{1-\gamma}{\rho}}l^{1-\gamma}}{1-\gamma} \left[1+o(1)\right].$$

Hence, we know that as  $x \to 0^+$ ,

$$\Delta R_t(\widetilde{x}) = \frac{(1-\alpha)^{\frac{1-\gamma}{\rho}} \left(\overline{L}_t^{1-\gamma} - L_t^{1-\gamma}\right)}{1-\gamma} \left[1 + o(1)\right] + (L_t - \overline{L}_t) w_t x$$
  
$$< -x$$
  
$$= -e^{\widetilde{x}}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough.

6. In the case of  $\rho < 0$  and  $\gamma = 1$ , (B.3) still holds. Similarly to the above, as  $x \to 0^+$ ,

$$U_{1}(c^{*}, l) - c^{*}x = \left[\frac{\log(1-\alpha)}{\rho} + \log l + o(1)\right]$$
$$-\left(\frac{\alpha}{1-\alpha}\right)^{\frac{1}{1-\rho}} l^{\frac{\rho}{\rho-1}} x^{1+\frac{1}{\rho-1}} [1+o(1)]$$
$$= \frac{\log(1-\alpha)}{\rho} + \log l + o(1),$$

and

$$\Delta R_t(\widetilde{x}) = \left[\log \overline{L}_t - \log L_t + o(1)\right] + (L_t - \overline{L}_t)w_t x$$
  
$$< -(1 + L_t w_t) x$$
  
$$= -(1 + L_t w_t)e^{\widetilde{x}}$$

provided x is enough near zero, i.e.,  $\tilde{x}$  is small enough. Then we have finished the proof.

**Lemma 10**  $\Delta R_t(\widetilde{x}) > 0$  for any  $\widetilde{x} > \log \overline{x}_t$ . Moreover, ess.inf  $\{\Delta R_t(\widetilde{x}) : (\omega, t) \in \Omega \times [0, T]\} > \epsilon e^{\widetilde{x}}$  provided  $\widetilde{x}$  is large enough, where  $\epsilon$  is a positive constant independent of  $\widetilde{x}$ .

**Proof** For any  $\tilde{x} > \log \overline{x}_t$ , we have that

$$\begin{aligned} \Delta R_t(\widetilde{x}) &= A_t(x) - \sup_{c \ge 0} \{ U_1(c, L_t) - xc \} + L_t w_t x \\ &= \sup_{c \ge 0, l \ge 0} \{ U_1(c, l) - x(c + w_t l) \} - \sup_{c \ge 0} \{ U_1(c, L_t) - xc \} + L w_t x \\ &> \sup_{c \ge 0} \{ U_1(c, L_t) - x(c + L_t w_t) \} - \sup_{c \ge 0} \{ U_1(c, L_t) - xc \} + L_t w_t x \\ &= 0. \end{aligned}$$

Moreover, if  $\tilde{x}$  is large enough, we know that

$$\Delta R_t(\widetilde{x}) \ge U_1(c_t^*, \overline{L}_t) - U_1(c_t^*, L_t) + (L_t - \overline{L}_t)w_t x$$

$$\ge (L_t - \overline{L}_t)x \left[ w_t - \frac{\partial_l U_1(c_t^*, \overline{L}_t)}{x} \right],$$
(B.4)

where  $c_t^* = \mathcal{J}_{U_1}(x; L_t)$ , and we have used the fact  $U_1$  is concave with respect to l. Applying the same method as in the proof of Lemma 9, we can show that as  $x \to +\infty$ ,

$$c_t^* = \begin{cases} \alpha^{\frac{1}{1-\rho}} (1-\alpha)^{\frac{1-\gamma-\rho}{\rho(1-\rho)}} L_t^{\frac{1-\gamma-\rho}{1-\rho}} x^{\frac{1}{\rho-1}} \Big[ 1+o(1) \Big], & 0 < \rho < 1; \\ \alpha^{\frac{1}{1-\alpha(1-\gamma)}} L_t^{\frac{(1-\alpha)(1-\gamma)}{1-\alpha(1-\gamma)}} x^{\frac{1}{\alpha(1-\gamma)-1}}, & \rho = 0; \\ \alpha^{\frac{1-\gamma}{\rho\gamma}} x^{\frac{-1}{\gamma}} \Big[ 1+o(1) \Big], & \rho < 0. \end{cases}$$

So we derive that as  $x \to +\infty$ ,

$$\frac{\partial_l U_1(c_t^*, \overline{L}_t)}{x} = \begin{cases} (1-\alpha) \overline{L}_t^{\rho-1} [\alpha(c_t^*)^{\rho} + (1-\alpha) \overline{L}_t^{\rho}]^{\frac{1-\gamma-\rho}{\rho}} x^{-1} \\ = o(1), & 1 > \rho \neq 0; \\ (1-\alpha) (c_t^*)^{\alpha(1-\gamma)} l^{(1-\alpha)(1-\gamma)-1} x^{-1} = o(1), & \rho = 0. \end{cases}$$

Hence, by (B.4), we conclude that  $\Delta R_t(\tilde{x}) > \epsilon e^{\tilde{x}}$  provided  $\tilde{x}$  is large enough, where  $\epsilon$  is a positive constant.

**Lemma 11** If  $\gamma + \rho \leq 1$ , then  $\partial_{\tilde{x}} \Delta R \geq 0$ .

**Proof** From the definition of  $\Delta R$ , we can show that

$$e^{-\widetilde{x}}\partial_{\widetilde{x}}\Delta R_{t}(\widetilde{x}) = \partial_{x}\widehat{U}_{1,t}(x) - \partial_{x}\widetilde{U}_{1,t}(x) + L_{t}w_{t}$$
  
$$= (-\mathcal{J}_{U_{1}}(x; l_{t}^{*}) - l_{t}^{*}w_{t}) - (-\mathcal{J}_{U_{1}}(x; L_{t})) + L_{t}w_{t}$$
  
$$= \mathcal{J}_{U_{1}}(x; L_{t}) - \mathcal{J}_{U_{1}}(x; l_{t}^{*}) + (L_{t} - l_{t}^{*})w_{t},$$
  
$$l_{t}^{*} = \overline{L}_{t}\min\left\{1, \left(\frac{\overline{x}_{t}}{x}\right)^{\frac{1}{\gamma}}\right\} \in (0, \overline{L}_{t}].$$

Moreover, it is easy to check that

$$\partial_{cl} U_1(c,l) = \begin{cases} (1-\gamma-\rho)\alpha(1-\alpha)c^{\rho-1}l^{\rho-1} \\ \times [\alpha c^{\rho} + (1-\alpha)l^{\rho}]^{\frac{1-\gamma-2\rho}{\rho}}, & 1 > \rho \neq 0; \\ \alpha(1-\alpha)(1-\gamma)c^{\alpha(1-\gamma)-1}l^{(1-\alpha)(1-\gamma)-1}, & \rho = 0. \end{cases}$$

Since  $1 - \gamma - \rho \ge 0$ , we have that  $\partial_c U_1(c, l)$  is increasing with respect to l, and  $\mathcal{J}_{U_1}(x; l)$  is increasing with respect to l by  $\partial_{cc} U_1(c, l) < 0$ . Hence,  $\partial_{\tilde{x}} \Delta R \ge 0$ .  $\Box$ 

## References

- Ball, C.A., Torous, W.N.: Stochastic correlation across international stock markets. J. Empir. Financ. 7(3–4), 373–388 (2000)
- Bensoussan, A., Friedman, A.: Nonlinear variational inequalities and differential games with stopping times. J. Funct. Anal. 16(3), 305–352 (1974)
- Bodie, Z., Merton, R.C., Samuelson, W.F.: Labor supply flexibility and portfolio choice in a life cycle model. J. Econ. Dyn. Control 16(3–4), 427–449 (1992)
- Choi, K.J., Shim, G.: Disutility, optimal retirement, and portfolio selection. Math. Financ. 16(2), 443– 467 (2006)
- Choi, K.J., Shim, G., Shin, Y.H.: Optimal portfolio, consumption-leisure and retirement choice problem with CES utility. Math. Financ. 18(3), 445–472 (2008)
- Cox, J.C., Huang, C.: Optimal consumption and portfolio policies when asset prices follow a diffusion process. J. Econ. Theory 49(1), 33–83 (1989)
- Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (1992)
- Du, K., Tang, S.: Strong solution of backward stochastic partial differential equations in C<sup>2</sup> domains. Prob. Theory Relat. Fields 154(1–2), 255–285 (2012)
- Dybvig, P.H., Liu, H.: Lifetime consumption and investment: retirement and constrained borrowing. J. Econ. Theory 145(3), 885–907 (2010)
- El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S., Quenez, M.C.: Reflected solutions of backward SDE's, and related obstacle problems for PDE's. Ann. Probab. 25(2), 702–C737 (1997)
- Farhi, E., Panageas, S.: Saving and investing for early retirement: a theoretical analysis. J. Financ. Econ. 83(1), 87–121 (2007)
- Friedman, A.: Stochastic games and variational inequalities. Arch. Ration. Mech. Anal. 51(5), 321–346 (1973)
- Heath, D., Jarrow, R., Morton, A.: Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica 60(1), 77–105 (1992)
- Hu, Y., Peng, S.: Adapted solution of a backward semilinear stochastic evolution equation. Stoch. Anal. Appl. 9, 445–459 (1991)
- Hu, Y., Ma, J., Yong, J.: On semi-linear degenerate backward stochastic partial differential equations. Probab. Theory Relat. Fields 123, 381–411 (2002)
- Inada, K.: On a two-sector model of economic growth: comments and a generalization. Rev. Econ. Stud. 30(2), 119–127 (1963)
- Karatzas, I., Lehoczky, J.P., Sethi, S.P., Shreve, S.E.: Explicit solution of a general consumption/investment problem. Math. Oper. Res. 11(2), 261–294 (1986)
- Karatzas, I., Lehoczky, J.P., Shreve, S.E.: Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. SIAM J. Control Optim. 25(6), 1557–1586 (1987)
- 19. Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance. Springer, New York (1998)
- Karatzas, I., Wang, H.: Utility maximization with discretionary stopping. SIAM J. Control Optim. 39(1), 306–329 (2000)
- Koo, H.K., Tang, S., Yang, Z.: A Dynkin game under Knightian uncertainty. Discret. Contin. Dyn. Syst. 35(11), 5467–5498 (2015)
- 22. Li, K., Liu, J.: Optimal dynamic momentum strategies. Working paper

- 23. Mao, X.: Stochastic Differential Equations and Applications. Woodhead Publishing, Cambridge (2011)
- Nutz, M., Zhang, J.: Optimal stopping under adverse nonlinear expectation and related games. Ann. Appl. Prob. 25(5), 2503–2534 (2015)
- Peskir, G., Shiryaev, A.: Optimal Stopping and Free-Boundary Problems. Birkhäuser Verlag, Basel (2006)
- Qiu, J., Wei, W.: On the quasi-linear reflected backward stochastic partial differential equations. J. Funct. Anal. 267(10), 3598–3656 (2014)
- Qiu, J.: Viscosity solutions of stochastic Hamilton-Jacobi-Bellman equations. SIAM J. Control Optim. 56(5), 3708–3730 (2018)
- Tang, S., Koo, H.K.: Real options: a framework of optimal switching. In: Koo, H.K. (ed.) New Frontiers of Financial Engineering. IOS Press, Amsterdam (2012)
- 29. Uzawa, H.: On a two-sector model of economic growth II. Rev. Econ. Stud. 30(2), 105–118 (1963)
- Yang, Z., Koo, H.K.: Optimal consumption and portfolio selection with early retirement option. Math. Oper. Res. 43(4), 1378–1404 (2018)
- Yang, Z., Tang, S.: Dynkin game of stochastic differential equations with random coefficients and associated backward stochastic partial differential variational inequality. SIAM J. Control Optim. 51(1), 64–95 (2013)
- 32. Yong, J., Zhou, X.Y.: Stochastic control: Hamiltonian systems and HJB equations. Springer, New York (1999)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.